

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
INSTITUTO DE INGENIERÍA



DISEÑO DE MÉTODO DE CALIBRACIÓN PARA EL AUMENTO
DE LA VERSATILIDAD EN LA INFORMACIÓN
ADQUIRIDA POR SISTEMA DE VISIÓN LÁSER Y
ESTEREOVISIÓN

TESIS
QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS
PRESENTA
FERNANDO LÓPEZ MEDINA

Director de tesis: Dr. Oleg Sergiyenko
Codirectora de tesis: Dra. Wendy Flores Fuentes

Resumen

Esta tesis presenta una metodología sistemática y parcialmente automatizada para la calibración intrínseca y extrínseca de un sistema híbrido que integra un Sistema de Visión Técnica (TVS, por sus siglas en inglés) con un sistema de estereovisión, constituyendo una etapa previa a su implementación en tareas como la reconstrucción tridimensional. La calibración extrínseca permite fusionar información geométrica desde distintas perspectivas, combinando la alta precisión del TVS con la densidad y textura visual aportadas por las cámaras estéreo.

El enfoque propuesto reduce la necesidad de utilizar patrones con geometría conocida, empleando en su lugar una superficie plana como elemento auxiliar. Esta superficie permite usar la estereovisión para extraer parámetros intrínsecos del TVS —como la inclinación y alineación de sus espejos, la relación grados/paso de los motores y el backlash mecánico—, los cuales se corrigen mediante ajustes mecánicos o compensaciones computacionales. Asimismo, posibilita ejecutar la calibración extrínseca del sistema combinado. Este planteamiento aprovecha la capacidad del TVS para proyectar un haz láser claramente identificable por las cámaras incluso en condiciones de baja iluminación, a diferencia de los patrones tradicionales.

La calibración extrínseca se lleva a cabo mediante escaneos simultáneos, estimando la transformación rígida entre nubes de puntos a través del ajuste de planos utilizando mínimos cuadrados y descomposición en valores singulares (SVD, por sus siglas en inglés), sin necesidad de marcadores ni algoritmos de optimización iterativa.

Los experimentos se realizaron bajo diferentes configuraciones de montaje. En una primera etapa, con los dispositivos montados sobre una mesa óptica, se obtuvieron errores de alineación entre puntos correspondientes inferiores a 1.5 cm en promedio. Más adelante, tras implementar ajustes en el sistema, la evaluación sobre un robot con movimiento rectilíneo intermitente arrojó errores del orden de 6 mm.

La principal contribución de este trabajo radica en la integración de la calibración intrínseca y extrínseca dentro de un procedimiento unificado que reduce la dependencia de patrones con marcas visibles, empleando en su lugar superficies planas y formulaciones algebraicas directas.

Abstract

This thesis presents a systematic and partially automated methodology for the intrinsic and extrinsic calibration of a hybrid system that integrates a Technical Vision System (TVS) with a stereo vision system. This process constitutes a preliminary stage prior to its implementation in tasks such as three-dimensional reconstruction. The extrinsic calibration enables the fusion of geometric information from different perspectives by combining the high precision of the TVS with the visual density and texture provided by stereo cameras.

The proposed approach minimizes the reliance on calibration targets with predefined geometry by instead utilizing a single planar surface as the auxiliary reference. This surface enables the use of stereo vision to extract intrinsic parameters of the TVS—such as the inclination and alignment of its mirrors, degrees-per-step ratio of the motors, and mechanical backlash—which are corrected through mechanical adjustments or computational compensations. Additionally, it enables the execution of the extrinsic calibration of the combined system. This approach takes advantage of the TVS’s ability to project a laser beam that can be clearly detected by the cameras even under low-light conditions, unlike traditional calibration patterns.

The extrinsic calibration is performed through simultaneous scans, estimating the rigid-body transformation between point clouds by fitting planes using least squares and singular value decomposition (SVD), without the need for fiducial markers or iterative optimization algorithms.

The experiments were conducted under different mounting configurations. In an initial stage, with the devices mounted on an optical table, alignment errors between corresponding points were found to be below 1.5 cm on average. Later, after implementing adjustments to the system, evaluation on a robot with intermittent rectilinear motion yielded errors on the order of 6 mm.

The main contribution of this work lies in the integration of intrinsic and extrinsic calibration into a unified procedure that reduces reliance on patterns with visible markers, relying instead on planar geometry and direct algebraic formulations.

Dedicatoria

A mi madre, por su apoyo incondicional
a lo largo de este camino.

Agradecimientos

Deseo expresar mi más sincero agradecimiento a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo financiero, mediante sus becas, que me ha permitido continuar con mis estudios de posgrado y desarrollar esta investigación.

A la Universidad Autónoma de Baja California (UABC), por brindarme la oportunidad de avanzar en mi formación académica y alcanzar el grado de Maestría en Ciencias.

Al Instituto de Ingeniería de la UABC, campus Mexicali, por proporcionarme las instalaciones y recursos necesarios para llevar a cabo este trabajo, así como por contar con docentes comprometidos, cuyo conocimiento ha sido clave en mi desarrollo como estudiante de posgrado.

A mi director de tesis, el Dr. Oleg Sergiyenko, por su liderazgo en este proyecto de investigación y por brindarme la oportunidad de participar en él, compartiendo generosamente su experiencia y conocimientos en el área.

A mi codirectora de tesis, la Dra. Wendy Flores, por su colaboración en este trabajo de investigación, así como por su valioso tiempo y orientación durante el proceso. Sus consejos han sido fundamentales en cada etapa del proyecto.

Finalmente, a mis compañeros, por su apoyo constante en las diversas actividades y desafíos que surgieron a lo largo de esta etapa académica. Sus ideas y aportaciones fueron clave para el progreso de este trabajo.

Tabla de Contenido

1. Introducción	1
1.1. Planteamiento del problema	3
1.2. Justificación	4
1.3. Objetivos	5
1.3.1. Objetivo general	5
1.3.2. Objetivos específicos	5
1.4. Hipótesis	6
1.5. Antecedentes	7
2. Marco teórico	15
2.1. Visión artificial y percepción multisensorial	15
2.1.1. Sistemas de visión basados en cámaras	16
2.1.2. Sistemas de estereovisión	17
2.1.3. Sistemas láser para visión artificial	18
2.1.4. Sistemas híbridos de visión artificial	18
2.1.5. Sistema híbrido Estéreo-TVS	20
2.1.6. Marcos de referencia en visión multisensorial	20
2.1.7. Calibración extrínseca en sistemas multisensoriales	21
2.2. Sistema de Visión Técnica (TVS)	22
2.2.1. Subsistema de posicionador	23
2.2.2. Subsistema de apertura	24
2.2.3. Triangulación dinámica y nube de puntos 3D	25
2.3. Calibración del sistema TVS	26
2.4. Calibración de la Cámara	27
2.4.1. Modelos de cámara	28
2.4.2. Transformación de coordenadas en la imagen	29
2.4.3. Matriz extrínseca: transformación del mundo a la cámara	30
2.4.4. Sensor de la cámara y conversión de coordenadas	32
2.4.5. Matriz intrínseca: conversión a píxeles	32
2.4.6. Matriz de proyección	33
2.4.7. Proceso de calibración experimental	33
2.5. Métodos de estimación de la matriz de proyección	34
2.5.1. Método DLT (Direct Linear Transform)	34
2.5.2. Método de Zhang	35
2.6. Modelado de distorsión óptica	35
2.6.1. Rectificación de imágenes	35
2.7. Calibración de un sistema de estereovisión	36
2.7.1. Geometría epipolar	38
2.7.2. Correspondencia estereoscópica	39
2.7.3. Matriz fundamental	40
2.7.4. Matriz esencial	41
2.7.5. Descomposición de la matriz esencial (rotación y traslación)	42
2.7.6. Calibración extrínseca de dos sistemas	42

3. Metodología	45
3.1. Procedimiento de calibración en el sistema híbrido Estéreo-TVS	45
3.1.1. Parámetros intrínsecos del TVS	46
3.1.2. Calibración extrínseca del sistema Estéreo-TVS	48
3.2. Calibración intrínseca del TVS	49
3.2.1. Inclinación individual de los espejos del sistema TVS	51
3.2.2. Alineación entre los espejos del posicionador y la apertura	54
3.2.3. Caracterización angular de los motores: relación grados por paso	55
3.2.4. Método propuesto para medición del backlash	57
3.3. Calibración extrínseca en mesa óptica	60
3.3.1. Adquisición de datos y generación de nubes de puntos	60
3.3.2. Traslación de nubes de puntos	62
3.3.3. Rotación mediante ajuste de planos	64
3.4. Ejemplo de aplicación: calibración extrínseca montando sistemas en robot con movimiento intermitente	70
4. Experimentación y resultados	73
4.1. Calibración intrínseca del sistema Estéreo-TVS	74
4.1.1. Calibración del sistema de estereovisión	77
4.1.2. Inclinación y alineación de los espejos del posicionador y la apertura	79
4.1.3. Caracterización experimental de la relación angular grados/paso	88
4.2. Calibración extrínseca del sistema Estéreo-TVS montado en mesa óptica	94
4.2.1. Configuración experimental para la calibración extrínseca	94
4.2.2. Adquisición de datos y construcción de nubes de puntos	96
4.2.3. Aplicación del método de transformación	98
4.2.4. Error de correspondencia posterior a la transformación	99
4.3. Calibración extrínseca del sistema Estéreo-TVS montado en robot con movimiento intermitente	103
5. Conclusiones	108
6. Apéndice A: Publicaciones	111
BIBLIOGRAFÍA	112

Índice de Tablas

4.1. Resumen estadístico de la relación grados/paso por dirección.	93
4.2. Factores y niveles empleados en diseño experimental para la evaluación del método de calibración extrínseca sobre mesa óptica.	96
4.3. Resumen de resultados de calibración extrínseca considerando 48 experimentos sobre mesa óptica.	101
4.4. Factores y niveles empleados en diseño experimental para la evaluación del método de calibración extrínseca sobre robot con movimiento intermitente.	104
4.5. Resumen de resultados de calibración extrínseca considerando 72 experimentos sobre robot con movimiento intermitente.	106

Índice de Figuras

2.1. Esquema de un sistema de estereovisión y su línea base.	17
2.2. Ejemplo esquemático de la superposición de campos de visión (FOV) de dos sensores heterogéneos montados en un sistema móvil.	19
2.3. Ejemplo de marcos de referencia global y local en un sistema multisensorial.	21
2.4. Esquema general del Sistema de Visión Técnica (TVS)	23
2.5. Componentes del subsistema posicionador del TVS	24
2.6. Componentes del subsistema apertura del TVS	25
2.7. Esquema del modelo de cámara pinhole y su proyección invertida.	29
2.8. Modelo geométrico de la proyección de un punto en el mundo a coordenadas de la cámara.	30
2.9. Ejemplo de objeto de calibración con un patrón visual. Este patrón permite establecer correspondencias entre puntos 3D en el mundo real y su proyección en la imagen.	34
2.10. Ejemplo del efecto de la distorsión radial sobre una imagen y su corrección mediante rectificación.	36
2.11. Triangulación con un sistema estereoscópico.	37
2.12. Geometría epipolar: relación entre puntos de imagen y líneas epipolares en visión estéreo.	39
2.13. Búsqueda de correspondencias en un sistema estéreo rectificado.	40
2.14. Clasificación de los métodos de calibración extrínseca entre sensores heterogéneos, adaptado de [21].	44
3.1. Diagrama general del método de calibración del sistema combinado Estéreo–TVS.	47
3.2. Patrón de ajedrez empleado en la calibración de cámaras.	49
3.3. Ubicación de la superficie, que se asume como plana, dentro del campo de visión compartido por el sistema estéreo y el TVS.	50
3.4. Comparación entre espejos inclinados y no inclinados.	53
3.5. Comparación entre espejos de apertura y posicionador coalineados.	55
3.6. Triángulos proyectados en el plano XY a partir de los puntos triangulados antes y después del desplazamiento del motor.	57
3.7. Procedimiento para la adquisición de datos y determinación de parámetros extrínsecos.	61
3.8. Comparativa entre las nubes de puntos generadas por los sistemas TVS y estéreo.	62
3.9. Representación esquemática del vector de traslación \mathbf{V}_t , definido como la diferencia entre los centroides de las nubes de puntos capturadas por el sistema estéreo y el sistema TVS.	64
3.10. Planos ajustados a partir de las nubes de puntos obtenidas por los sistemas TVS y estéreo.	66
3.11. Comparación entre las nubes de puntos del sistema estéreo y TVS, incluyendo planos ajustados, centroides y vectores ortonormales definidos.	67
3.12. Resultado final del alineamiento entre las nubes de puntos TVS y estéreo tras la aplicación de transformaciones geométricas.	69
3.13. Esquema metodológico para la estimación de parámetros extrínsecos en un sistema Estéreo–TVS montado sobre un robot móvil.	71
4.1. Prototipo del Sistema de Visión Técnica (TVS) llamado "Mini TVS".	74
4.2. Módulo de estereovisión GXIVision con dos cámaras LSM22100.	75

4.3. Configuración experimental con el TVS, el sistema estéreo y el patrón plano montados sobre una mesa óptica.	76
4.4. Sistema de referencia asociado a la mesa óptica y al TVS, empleado como marco global para la interpretación de resultados.	76
4.5. Tablero de ajedrez utilizado en la fase inicial del método desarrollado para la calibración del sistema combinado Estéreo-TV S.	77
4.6. Ubicación del láser puntual sobre el espejo de la apertura para su calibración de inclinación.	80
4.7. Ejemplo de resultado en la evaluación de la inclinación en los espejos del sistema TVS. Se observa el ajuste lineal del haz verde, así como la detección del centroide de la mancha del láser infrarrojo.	82
4.8. Distribución de centroides y líneas ajustadas por imagen en experimento realizado antes del ajuste.	83
4.9. Evaluación de inclinación y alineación de espejos previa al ajuste, vista izquierda. . .	84
4.10. Evaluación de inclinación alineación de espejos previa al ajuste, vista derecha. . . .	84
4.11. Ejemplo representativo de reducción correcta de distancia entre láseres, posterior al ajuste.	85
4.12. Distribución de centroides y líneas ajustadas por imagen en experimento realizado después del ajuste.	86
4.13. Evaluación de inclinación y alineación de espejos después del ajuste, vista izquierda. .	86
4.14. Evaluación de inclinación y alineación de espejos después del ajuste, vista derecha. .	87
4.15. Visualización tridimensional de los puntos triangulados en un experimento de desplazamiento hacia la izquierda.	89
4.16. Cálculo angular en el plano XY para un experimento individual.	90
4.17. Distribución de la relación grados/paso en la dirección izquierda.	91
4.18. Distribución de la relación grados/paso en la dirección derecha.	91
4.19. Distribución de la relación grados/paso en la dirección arriba.	92
4.20. Distribución de la relación grados/paso en la dirección abajo.	92
4.21. Disposición del equipo en la mesa óptica para la calibración extrínseca.	95
4.22. Ubicaciones del patrón en la mesa óptica durante la experimentación.	95
4.23. Detección del contorno y centroide de la mancha láser bajo diferentes condiciones de iluminación. a): luz ambiental, b): oscuridad.	97
4.24. Ejemplo de nubes de puntos obtenidas en la experimentación.	98
4.25. Nube de puntos original del TVS y nube de puntos transformada del sistema de estereovisión con planos ajustados y vectores de referencia.	99
4.26. Error de correspondencia en el eje X en experimentos sobre mesa óptica.	100
4.27. Error de correspondencia en el eje Y en experimentos sobre mesa óptica.	101
4.28. Error de correspondencia en el eje Z en experimentos sobre mesa óptica.	101
4.29. Configuración experimental: robot móvil con sistema combinado Estéreo-TV S, junto con la computadora utilizada para el control del robot y el osciloscopio Tektronix TBS2000 para el monitoreo de la señal.	103
4.30. Error de correspondencia en el eje X para luz ambiental y oscuridad, representando en cada tono de gris una distancia a la superficie: 20 <i>cm</i> , 30 <i>cm</i> y 40 <i>cm</i>	104
4.31. Error de correspondencia en el eje Y para luz ambiental y oscuridad, representando en cada tono de gris una distancia a la superficie: 20 <i>cm</i> , 30 <i>cm</i> y 40 <i>cm</i>	105
4.32. Error de correspondencia en el eje Z para luz ambiental y oscuridad, representando en cada tono de gris una distancia a la superficie: 20 <i>cm</i> , 30 <i>cm</i> y 40 <i>cm</i>	105

Capítulo 1: Introducción

Los sistemas de visión artificial constituyen herramientas fundamentales en una amplia gama de aplicaciones industriales y científicas, donde su capacidad para adquirir y procesar información tridimensional del entorno resulta crucial para una interpretación precisa de la escena. Gracias a su versatilidad, estos sistemas se emplean exitosamente en tareas como navegación autónoma de vehículos en entornos urbanos [1], inspección de superficies [2], detección de fallas estructurales en tuberías [3], operaciones de rescate [4] y reconstrucción tridimensional de objetos [5].

Existen diversos tipos de sensores de visión artificial capaces de generar representaciones tridimensionales, cada uno con principios de funcionamiento, fortalezas y limitaciones particulares. Comprender estas diferencias resulta esencial para seleccionar el sistema más adecuado según los requerimientos específicos de la aplicación. En las siguientes secciones se describen los sistemas de visión basados en cámaras, así como los sistemas activos basados en láser.

Sistemas de visión con cámaras

Los sistemas de visión con cámaras se pueden clasificar en cuatro categorías principales: monoscópicos, estereoscópicos, visión omnidireccional y multivisión.

- **Monoscópicos:** Utilizados ampliamente por su bajo costo y simplicidad, permiten detectar objetos, clasificarlos y segmentar imágenes [6]. No obstante, presentan limitaciones en la estimación precisa de profundidad, al depender de modelos geométricos o supuestos de escala.
- **Estereoscópicos:** Utilizan dos cámaras calibradas para estimar la profundidad mediante triangulación, calculando mapas de disparidad entre imágenes correspondientes. Son útiles para reconstrucción 3D en tiempo real [7], pero su desempeño se degrada en regiones con baja textura o condiciones lumínicas adversas.
- **Visión omnidireccional:** Estos sistemas permiten capturar imágenes con campo visual de

360° en el plano horizontal y hasta 180° en elevación, utilizando configuraciones como lentes *fisheye* (dioptricas), espejos curvos (catadióptricas) o arreglos multicámara (polidióptricas) [8].

Los sistemas catadióptricos centrales cumplen con la propiedad de punto de vista único, lo que permite aplicar geometría epipolar y reconstrucción 3D. No obstante, requieren modelos de proyección más complejos y son sensibles a aberraciones ópticas y errores de alineación.

- **Multivisión:** Emplean múltiples cámaras dispuestas en distintas posiciones para aumentar cobertura visual y mejorar precisión en la reconstrucción tridimensional [9]. Aunque potentes, estos sistemas implican mayor complejidad en sincronización, procesamiento y calibración extrínseca.

Sistemas de escaneo láser y TVS

Los sistemas activos como el escaneo láser han cobrado gran relevancia en metrología, navegación y monitoreo estructural [10]. Técnicas como Time-of-Flight (ToF) [11] y Triangulación Dinámica (TD) [12] permiten obtener mediciones de alta precisión incluso en condiciones de poca iluminación. Sin embargo, presentan limitaciones asociadas a su costo, sensibilidad a ciertas superficies, y ausencia de información visual como color o textura [13].

Dentro de esta categoría, el LiDAR (Light Detection and Ranging) destaca como tecnología basada en el principio de ToF, generando nubes de puntos que representan la geometría del entorno. En contraste, el Sistema de Visión Técnica (TVS) opera mediante triangulación dinámica, calculando coordenadas tridimensionales a partir del ángulo de emisión, reflexión y geometría interna del sistema [12]. Su uso ha sido reportado en aplicaciones como navegación autónoma y análisis estructural [10].

Fusión de sensores y calibración en sistemas de visión

La integración de sensores heterogéneos es cada vez más común, ya que permite compensar las limitaciones individuales de cada sistema y aprovechar sus fortalezas combinadas. Para lograr una fusión efectiva, es imprescindible que los sensores compartan un marco de referencia común, lo cual

requiere procesos de calibración intrínseca y extrínseca.

Este trabajo se enfoca en un sistema híbrido Estéreo–TVS, que combina la densidad espacial de un sistema estéreo con la precisión puntual del TVS. Se propone un método de calibración robusto que permite estimar los parámetros intrínsecos del TVS y la transformación extrínseca entre ambos sensores sin recurrir a patrones físicos con marcas características y reduciendo la intervención manual, permitiendo operar incluso en condiciones de oscuridad. Esta metodología aprovecha la información tridimensional proporcionada por la estereovisión como herramienta auxiliar para estimar parámetros críticos del TVS.

1.1. Planteamiento del problema

En el laboratorio de Optoelectrónica del Instituto de Ingeniería de la Universidad Autónoma de Baja California (UABC) se cuenta con un prototipo que integra dos sistemas de visión artificial: un Sistema de Visión Técnica (TVS) y un sistema de estereovisión basado en cámaras. Aunque esta configuración ofrece un gran potencial para tareas de reconstrucción tridimensional y análisis geométrico, actualmente no se dispone de un procedimiento formal y validado que permita calibrar ambos sistemas de forma conjunta y precisa.

En particular, el TVS utilizado en este prototipo incorpora motores a pasos para el posicionamiento angular del haz láser mediante una transmisión mecánica. Sin embargo, no existía hasta ahora un método estandarizado para determinar parámetros intrínsecos fundamentales para el cálculo de coordenadas tridimensionales, como la relación grados/paso del motor, el backlash inducido por la transmisión, o la alineación entre los componentes del sistema. La ausencia de esta calibración sistemática introduce errores que afectan la precisión del cálculo de coordenadas del sistema.

Por otro lado, la calibración extrínseca entre el sistema estéreo y el TVS, es decir, la determinación de la transformación rígida que relaciona espacialmente ambos sistemas, se realizaba de forma manual o empírica. Frecuentemente se asumía únicamente una traslación entre ambos sistemas, omitiendo las componentes rotacionales. Esta simplificación compromete la coherencia geométrica de los datos fusionados, afectando directamente la calidad de las reconstrucciones tridimensionales.

Por tanto, se requiere una metodología sistemática y precisa para la calibración intrínseca del TVS y su alineación extrínseca con el sistema de estereovisión, que permita operar con fiabilidad bajo condiciones adversas y garantice la compatibilidad geométrica de los datos capturados por ambos sistemas.

1.2. Justificación

La precisión en los sistemas de visión por computadora es fundamental en aplicaciones como la navegación autónoma, la detección de fallas estructurales y la reconstrucción tridimensional de superficies. Para garantizar mediciones confiables y replicables, se requiere una calibración rigurosa que minimice errores geométricos y asegure la coherencia espacial de la información adquirida [14].

En el caso de sensores ópticos como las cámaras, la calibración intrínseca permite modelar la proyección entre el mundo tridimensional y el plano de imagen, considerando parámetros como la distancia focal, el centro óptico y las distorsiones. En configuraciones multisensoriales, es además necesaria la calibración extrínseca, que define las rotaciones y traslaciones entre sistemas independientes, siendo indispensable para fusionar datos en un marco de referencia común.

El sistema Estéreo-TV S abordado en este trabajo requiere tres etapas clave de calibración: calibración de las cámaras (intrínseca y extrínseca), intrínseca del TV S, y extrínseca entre ambos. Si bien se emplea el método clásico de Zhang [15] para la calibración inicial de las cámaras, este patrón no se requiere en etapas posteriores. Esta independencia de patrones físicos con marcas características permite realizar la calibración incluso en completa oscuridad, superando limitaciones tradicionales.

La metodología desarrollada contempla lo siguiente:

- **Calibración del sistema estéreo:** Estimación de parámetros intrínsecos y extrínsecos entre cámaras. Estos datos permiten, posteriormente, emplear el sistema estéreo como herramienta auxiliar para la calibración del TV S.
- **Calibración intrínseca del TV S:** Se estiman parámetros como la relación grados/paso, alineación óptica y compensación del backlash, empleando visión estéreo como referencia

auxiliar para cuantificar estos valores sin medición manual.

- **Calibración extrínseca:** Se proyecta una mancha láser sobre superficies planas observadas por ambos sistemas, generando nubes de puntos con las cuales se ajustan planos y se calcula la transformación entre sistemas.

Hasta este trabajo, no se había documentado en la literatura una metodología formal para la calibración conjunta entre un TVS y un sistema de cámaras. Además, este trabajo plantea dos enfoques complementarios: (1) calibrar el TVS utilizando información de las cámaras (enfoque principal), y (2) emplear el TVS como sistema de referencia para calibrar las cámaras (propuesto como línea futura). La eventual implementación de un procedimiento iterativo entre ambos sistemas permitiría reducir aún más la dependencia de patrones visuales externos y aumentar la autonomía del sistema en escenarios operativos reales.

1.3. Objetivos

1.3.1. Objetivo general

Desarrollar un método sistemático de calibración intrínseca y extrínseca para un sistema híbrido de visión artificial compuesto por un sistema de visión láser y un sistema de estereovisión, con el fin de mejorar la precisión geométrica, la coherencia espacial y la robustez de las mediciones tridimensionales, incluso en condiciones adversas.

1.3.2. Objetivos específicos

- Determinar con precisión los parámetros intrínsecos del sistema de estereovisión mediante técnicas estándar basadas en patrones de calibración, garantizando una correcta reconstrucción tridimensional por triangulación.
- Utilizar la información obtenida por el sistema de estereovisión para identificar, evaluar y ajustar aquellos parámetros del TVS implicados directamente en el cálculo de coordenadas espaciales, asegurando su estabilidad y exactitud durante la operación del sistema.

- Desarrollar un método robusto y automatizado basado en la proyección de un láser sobre superficies planas para realizar la calibración extrínseca entre el sistema TVS y el sistema de estereovisión, generando algoritmos capaces de determinar automáticamente las transformaciones geométricas (rotaciones y traslaciones) que alinean ambos sistemas dentro de un marco de referencia global común.
- Diseñar un procedimiento específico para calibrar el sistema combinado Estereo-TVS en condiciones adversas, tales como baja iluminación o total oscuridad, aprovechando la robustez de la detección de manchas láser proyectadas sobre las superficies analizadas.
- Identificar las limitaciones técnicas y operativas del método propuesto, formulando recomendaciones claras para su mejora y estableciendo líneas futuras de investigación que permitan ampliar su ámbito de aplicación.
- Incorporar en la interfaz de usuario existente del TVS las herramientas necesarias para ejecutar los procedimientos de calibración propuestos, facilitando así la interacción del operador con el sistema y optimizando la gestión y control de los algoritmos implementados.

1.4. Hipótesis

Es posible desarrollar un método integral de calibración para el sistema híbrido Estereo-TVS utilizando el sistema de estereovisión como herramienta auxiliar para identificar, determinar y ajustar los parámetros intrínsecos del TVS necesarios para obtener mediciones precisas de coordenadas tridimensionales. Asimismo, se plantea que mediante la captura de imágenes de las proyecciones láser generadas por el TVS, utilizando el sistema de estereovisión, se obtendrá una nube de puntos homólogos que permitirá realizar una calibración extrínseca efectiva entre ambos sistemas. Dicho método será capaz de operar incluso en condiciones adversas de iluminación, facilitando la integración robusta y precisa de la información proveniente de ambos sensores en una única nube de puntos referenciada a un marco global común.

1.5. Antecedentes

Existen diversos sensores empleados en sistemas de percepción para navegación autónoma, entre los que se incluyen cámaras RGB o multiespectrales, escáneres LiDAR, telémetros láser (LRF, por sus siglas en inglés), así como sensores no ópticos como el ultrasonido y el radar. En este contexto, la visión artificial cumple un rol central al proporcionar información estructural del entorno basada en datos visuales. Una de las etapas fundamentales para asegurar la precisión de estos sistemas es la calibración, proceso que comprende tanto la determinación de parámetros intrínsecos como extrínsecos.

El sistema de visión técnica (TVS) constituye una solución activa para la reconstrucción tridimensional mediante escaneo láser, basada en el principio de triangulación dinámica. A diferencia de los sensores pasivos tradicionales, como las cámaras RGB o sistemas estéreo, el TVS proyecta un haz láser y, mediante la medición de ángulos relativos entre la fuente emisora, el punto de reflexión y la apertura de escaneo, calcula con alta precisión las coordenadas tridimensionales del entorno [12]. Esta técnica permite estimar directamente la posición 3D de cada punto sin requerir emparejamiento estéreo ni reconstrucción densa por disparidad, y además presenta la ventaja de operar con fiabilidad en condiciones de baja iluminación o textura deficiente.

El diseño del TVS involucra un sistema óptico-mecánico conformado por una fuente láser, una apertura de escaneo y un conjunto de sensores fotodetectores sincronizados mediante controladores eléctricos. Este conjunto permite la construcción de nubes de puntos con propiedades metrológicas, lo que lo hace especialmente adecuado para tareas de navegación autónoma, mapeo y monitoreo estructural. En [12] se demuestra su implementación en robots móviles, destacando su superioridad frente a métodos ópticos que se basan en estimaciones probabilísticas en lugar de mediciones espaciales directas.

Una línea de investigación emergente se enfoca en optimizar el rendimiento del TVS, particularmente en lo relativo a su velocidad de escaneo. En [16] se propone una técnica de paso angular variable combinado (combined variable step), que ajusta dinámicamente la resolución del escaneo

en función de la complejidad geométrica del entorno. Mediante simulaciones y redes neuronales, se identifican configuraciones angulares óptimas que permiten reducir significativamente el tiempo de adquisición sin sacrificar la calidad en zonas críticas.

En escenarios colaborativos, como los sistemas de robótica en enjambre, el TVS se ha integrado con algoritmos de gestión inteligente de datos. En [17] se presenta una arquitectura distribuida donde múltiples TVS montados en robots cooperativos escanean simultáneamente distintas secciones de un entorno, generando un mapa 3D unificado. Este sistema, apoyado en técnicas de inteligencia artificial y almacenamiento distribuido, mejora el rendimiento de tareas como el dead-reckoning (estimación de posición basada en desplazamientos previos) y la planificación de trayectorias en entornos dinámicos y no estructurados.

La diagnosis de fallas en los componentes del TVS también ha sido objeto de estudio. En particular, se ha demostrado que pequeñas inestabilidades en la velocidad de rotación del motor de escaneo pueden inducir errores sistemáticos en la estimación de las coordenadas. Para mitigar estos efectos, se han desarrollado observadores en modo deslizante (Sliding Mode Observers, SMO) capaces de detectar variaciones inadmisibles y descartar datos afectados antes de su procesamiento, mejorando la fiabilidad del sistema [18].

Si bien el TVS ha mostrado aplicaciones industriales y móviles, su versatilidad ha permitido extenderlo incluso al ámbito médico. En [19] se implementa un TVS adaptado para la evaluación tridimensional no invasiva de deformidades musculoesqueléticas, como la escoliosis. Gracias a su principio activo y a la independencia de condiciones lumínicas externas, el sistema permite la reconstrucción precisa de la superficie corporal sin recurrir a radiación ionizante ni contacto físico.

A pesar de estas contribuciones, la calibración del TVS sigue siendo una línea de investigación poco desarrollada. La mayoría de los trabajos se concentran en su caracterización geométrica o en la mejora de su arquitectura, sin establecer protocolos sistemáticos para su calibración intrínseca y extrínseca. Esta laguna representa una oportunidad para explorar metodologías robustas de alineación en escenarios donde el TVS se integra con otros sensores, particularmente en configuraciones heterogéneas donde la precisión de la fusión sensorial depende críticamente del conocimiento exacto

de las transformaciones de cuerpo rígido involucradas.

No obstante, la correcta estimación de estas transformaciones, específicamente la calibración extrínseca, representa un desafío recurrente en configuraciones multisensor heterogéneas. Esta problemática ha motivado el desarrollo de diversas estrategias de alineación espacial entre sensores, cuyas soluciones abarcan desde métodos analíticos con solución explícita (closed-form), que pueden resolverse directamente mediante fórmulas algebraicas, hasta esquemas de optimización iterativa que presentan funciones objetivo no convexas, es decir, con múltiples mínimos locales que dificultan encontrar la solución óptima. Estas metodologías pueden clasificarse en función del tipo de correspondencias geométricas que emplean (punto-a-plano, plano-a-plano, línea-a-plano, entre otras), del nivel de intervención humana requerido y de su dependencia en estructuras físicas externas como patrones impresos o dispositivos inerciales auxiliares.

En este contexto, la presente tesis se sitúa en una línea de trabajo que busca maximizar la observabilidad estructural del sistema de fusión sensorial utilizando únicamente elementos del entorno, evitando toda dependencia de infraestructura adicional. La estrategia propuesta se inscribe dentro de un enfoque de calibración explícitamente pasivo en cuanto a patrones, pero activo en cuanto a adquisición láser, y se basa en la extracción de características geométricas implícitas en superficies planas. Esta elección metodológica conlleva ventajas significativas en términos de portabilidad operativa, robustez ante perturbaciones mecánicas y adaptabilidad espacial a diferentes entornos no controlados.

En las siguientes secciones se abordarán distintos enfoques reportados en la literatura para resolver el problema de la calibración extrínseca entre sensores disímiles, contrastando sus requisitos técnicos, su aplicabilidad en escenarios reales y sus limitaciones. Esta revisión permitirá delimitar con precisión los vacíos existentes en el estado del arte y justificar la pertinencia de una propuesta basada en la correspondencia plano-a-plano (para rotación) y punto-a-punto (traslación), sin marcadores ni condiciones de contorno, aplicando soluciones algebraicas directas como el ajuste de planos mediante SVD.

Comparación de estrategias de calibración extrínseca

En [20], la calibración extrínseca entre el sistema estéreo y el TVS se realiza mediante una estrategia basada directamente en las dimensiones físicas del prototipo que ha sido construido mediante manufactura aditiva, también conocida como impresión 3D. Se define entonces una transformación rígida homogénea que relaciona el marco de coordenadas del sistema estéreo con el Sistema de Visión Técnica por Escaneo Láser (LSTVS, por sus siglas en inglés). Dicha transformación está compuesta por una matriz de rotación $R(\theta, \varphi, \psi)$, que depende de los ángulos de Euler (yaw, pitch y roll), y un vector de traslación $t = (t_x, t_y, t_z)$, el cual describe el desplazamiento entre los orígenes de ambos sistemas. Estos parámetros fueron deducidos directamente del diseño CAD del prototipo. Por ejemplo, para el sistema de cámaras de definición estándar (SD, por sus siglas en inglés), se aplica una rotación de -7° en el eje X y una traslación de 12.5 cm en Z . Este enfoque presupone una estructura rígida perfectamente conocida y estable, sin mecanismos de verificación o corrección ante desviaciones geométricas inducidas por tolerancias de ensamblaje o desajustes temporales.

En contraste, el método propuesto en esta tesis emplea una estrategia donde la transformación extrínseca se estima a partir de una superficie plana, externa al sistema, que es observada simultáneamente por ambos sistemas. Mediante las nubes de puntos obtenidas al escanear esta superficie se logra estimar la traslación y rotación. A diferencia del método previo, esta propuesta es adecuada para sistemas que necesitan recalibración periódica, y ofrece mayor robustez frente a incertidumbres geométricas o variabilidad mecánica.

Pese a la limitada trayectoria del TVS en la bibliografía, pueden trazarse paralelismos con la evolución de la calibración extrínseca en otros dispositivos de escaneo láser combinados con cámaras. En este ámbito, se han desarrollado diversos métodos que abarcan el uso de patrones físicos de referencia y algoritmos avanzados de optimización no lineal para la estimación de parámetros extrínsecos. En las siguientes secciones, se discutirán ejemplos representativos para ilustrar sus rasgos distintivos, resaltar su similitud con nuestro método y evidenciar la forma en que nuestra aproximación aborda vacíos existentes, principalmente en lo relativo a la dependencia de patrones con

marcas visibles y a la capacidad del método para operar de forma fiable incluso en condiciones de baja iluminación.

A modo de comparación, se describirán varios métodos de calibración extrínseca que integran un Laser Range Finder con cámaras, dispositivos muy utilizados en el mapeo bidimensional o tridimensional, la navegación y la detección de obstáculos. La calibración extrínseca en estos escenarios resulta esencial, ya que proporciona los parámetros de rotación y traslación requeridos para la superposición correcta de datos entre sensores [21]. Dichos métodos pueden clasificarse en aquellos que requieren el uso de patrones de referencia y aquellos que no, así como en enfoques manuales y en procedimientos automatizados. Esta categorización brinda una visión panorámica de las diversas estrategias metodológicas y, simultáneamente, subraya la necesidad de esquemas de calibración más robustos y versátiles, como el que aquí se propone.

Existen métodos que encuentran correspondencias entre punto a punto, punto a línea, punto a plano, línea a plano, plano a plano, basados en movimiento y basados en dependencia. Nuestro método encuentra correspondencias plano a plano (para rotación) y punto a punto (para traslación), ya que ambos sistemas escanean una superficie plana cualquiera y encuentra la transformación necesaria para alinearlos.

En los últimos años, la calibración extrínseca entre sensores heterogéneos se ha convertido en un área crítica dentro de los sistemas de percepción para plataformas móviles y aplicaciones de visión artificial. Diversos métodos han sido propuestos para estimar la transformación rígida entre sensores, siendo los enfoques manuales los más comunes en etapas tempranas, donde el usuario selecciona características compartidas en ambos dominios sensoriales. Sin embargo, esta intervención humana introduce errores sistemáticos y limita la repetibilidad del procedimiento. Para mitigar estas limitaciones, han surgido métodos automáticos, los cuales se apoyan en objetos de calibración con geometría conocida o en superficies de referencia que permiten la extracción robusta de características geométricas (como esquinas, bordes o planos) sin intervención del operador.

Tradicionalmente, estos métodos recurren a patrones estructurados, sensores auxiliares o procedimientos de optimización iterativa. En contraste, el método que aquí se propone elimina varias de

estas dependencias, utilizando únicamente superficies planas naturales, sin necesidad de bordes visibles, sensores inerciales ni estructuras geométricas impresas. A continuación, se realiza un análisis comparativo frente a metodologías representativas que ejemplifican las limitaciones que este trabajo resuelve.

En [22] se propone un método basado en la observación conjunta de un tablero de ajedrez para calibración por una cámara y un LRF 2D. Aunque presentan una solución cerrada inicial, requieren una posterior optimización no lineal para minimizar el error de reproyección. Además, su modelo depende de múltiples vistas del patrón y de la correcta segmentación de la intersección entre el plano láser y el patrón visual. El método aquí propuesto prescinde de cualquier patrón marcado y no requiere múltiples adquisiciones, gracias al uso de correspondencias punto a punto capturadas simultáneamente.

En [23] se extiende este tipo de calibración al caso de cámaras omnidireccionales, usando proyecciones unificadas. A pesar de su generalidad, el enfoque sigue dependiendo de patrones visibles y estructuras visuales impresas, y no es robusto a condiciones de baja iluminación. Por el contrario, la propuesta actual utiliza la alta intensidad lumínica del láser como ancla visual, detectable incluso en condiciones de escasa iluminación.

En [24] se introduce un objetivo (target) triplanar con múltiples checkerboards visibles desde distintos ángulos. Este enfoque mejora la estabilidad numérica en estimaciones desde una sola captura, pero incrementa significativamente la complejidad del hardware requerido. En cambio, el presente método elimina la necesidad de patrones físicos, capturando planos estructurales presentes en el entorno.

En [25] se emplea un patrón triangular especializado para generar restricciones tipo línea-a-plano. Si bien obtienen precisión mediante correspondencias geométricas bien definidas, dependen de la correcta detección de bordes y orientación angular específica del patrón, lo que impone condiciones de captura muy controladas. La solución propuesta evita tales dependencias gracias al uso de planos sin requerimientos sobre su contorno visible ni orientación específica.

En [26] se propone una solución algebraica para correspondencias plano-línea, formulada como

un sistema polinomial. Su ventaja radica en la estabilidad numérica, pero sigue requiriendo líneas detectables y visibles en imagen, lo cual es inviable en entornos oscuros o con texturas pobres. Nuestra propuesta, al utilizar directamente la mancha proyectada del láser, asegura correspondencias robustas independientemente de las texturas del entorno.

En [27] se presenta un enfoque cerrado que emplea un patrón trirectangular como referencia geométrica, resolviendo la calibración extrínseca mediante un problema de *perspectiva de tres puntos* (P3P, por sus siglas en inglés) para el LRF y un problema de *perspectiva de tres líneas* (P3L, por sus siglas en inglés) para la cámara. Aunque esta formulación permite una solución algebraica sin recurrir a optimización iterativa, requiere una estructura física rígida y ortogonal, lo cual limita su uso a entornos controlados. En contraste, nuestro método se basa en la detección de planos arbitrarios sin marcas visibles y en el uso de correspondencias directas entre nubes de puntos, lo que permite una mayor adaptabilidad en escenarios no estructurados.

En [28] se propone una aproximación más cercana a la presentada en esta tesis, al utilizar un target plano sin marcas visuales. Sin embargo, su método sigue dependiendo de restricciones geométricas como líneas de borde, y resuelve las correspondencias mediante optimización no lineal. Nuestro método, en cambio, utiliza directamente todos los puntos de las nubes capturadas, y aplica ajuste por SVD para extraer planos sin requerir bordes explícitos.

En [29] y [30] se incorporan datos inerciales (IMU) para estimar la rotación entre sensores, reduciendo el número de correspondencias necesarias. Aunque esto mejora la eficiencia algorítmica, introduce dependencia de hardware adicional. La presente propuesta logra una calibración completa sin necesidad de sensores adicionales, utilizando exclusivamente la geometría implícita en la captura simultánea de la mancha láser.

En [31] se utiliza un patrón circular para la calibración entre una cámara y un LiDAR multilámina, aprovechando su simetría para estimar de forma más estable los centros de los círculos detectados en imagen y nube de puntos. Esta configuración mitiga el impacto del ruido, a diferencia de los métodos que dependen de esquinas o bordes. Sin embargo, la necesidad de un patrón físico específico y la variación de resolución entre capas del LiDAR limitan su aplicabilidad. En contraste, nuestro

método no requiere targets especializados ni depende de sensores multilámina.

En [30] se utiliza un punto brillante (láser visible) como referencia móvil, capturado por un sistema estéreo y un LIDAR 3D, combinado con datos de una IMU para obtener correspondencias 3D. Aunque el uso del punto brillante es análogo a nuestra captura de la mancha del TVS, su dependencia de sensores inerciales y el uso de múltiples planos inclinados en la adquisición agregan complejidad innecesaria para escenarios fijos.

Finalmente, en [32] se propone un sistema de calibración online mediante optimización continua de parámetros extrínsecos, combinando estimación de profundidad monocular, detección de bordes y registros sucesivos a través del algoritmo Iterative Closest Point (ICP, por sus siglas en inglés). Este enfoque, diseñado para vehículos autónomos en movimiento, ofrece una solución robusta y flexible, aunque con un costo computacional elevado y dependencia de condiciones de navegación continua. En contraste, nuestro método está orientado a escenarios estáticos o con desplazamiento intermitente, y por tanto, constituye una alternativa más simple, determinista y de bajo costo para tareas de calibración local.

Capítulo 2: Marco teórico

2.1. Visión artificial y percepción multisensorial

La visión artificial es una disciplina multidisciplinaria que integra fundamentos del procesamiento digital de imágenes, geometría computacional, óptica física y aprendizaje automático. Su objetivo es dotar a los sistemas artificiales de capacidades perceptuales que emulan, en parte, las funciones del sistema visual humano, mediante la adquisición, interpretación y análisis automático de información visual proveniente del entorno. Esta área ha evolucionado significativamente, con aplicaciones consolidadas en la industria, la biometría, la robótica y la astronomía, consolidándose como un campo con proyección tanto científica como tecnológica [33].

Estos sistemas combinan sensores ópticos pasivos, como cámaras monoculares o estéreo, que capturan radiación electromagnética reflejada sin emitir señales propias, con sensores ópticos activos, como escáneres LiDAR o dispositivos basados en triangulación láser, que proyectan luz estructurada o pulsada para inferir la geometría tridimensional del entorno. Mientras los sensores activos requieren fuentes de iluminación artificial y suelen incluir cámaras de tiempo de vuelo (ToF) o escáneres láser, los pasivos operan en función de la iluminación ambiental disponible. Esta combinación permite una percepción tridimensional robusta del entorno, optimizando la resolución espacial, la precisión métrica y la tolerancia frente a condiciones adversas como baja iluminación o presencia de partículas suspendidas en el aire [34].

La elección del tipo de sensor depende de la aplicación concreta. Por ejemplo, los sensores láser permiten la reconstrucción tridimensional precisa del entorno mediante mediciones directas de distancia y geometría a través de técnicas como el tiempo de vuelo (ToF). Estas capacidades resultan esenciales en aplicaciones como cartografía digital, modelado estructural, navegación autónoma en interiores y metrología industrial [35]. Las cámaras ópticas, incluyendo sensores RGB, infrarrojos (IR) y multiespectrales como los SWIR (Short-Wave Infrared), capturan información radiométrica,

es decir, miden la intensidad de la radiación electromagnética reflejada o emitida por las superficies en distintas bandas espectrales. Esta información no solo permite reconstruir la apariencia visual, sino también inferir propiedades físicas o estructurales del material observado. Gracias a su alta resolución espacial y sensibilidad espectral, estos sensores son especialmente útiles en tareas de inspección superficial, detección de defectos, clasificación de materiales y reconocimiento de patrones, incluso en condiciones donde la percepción visual convencional es limitada [36].

2.1.1. Sistemas de visión basados en cámaras

Los sistemas de visión por computadora permiten el análisis automatizado de imágenes adquiridas por sensores ópticos, usualmente cámaras digitales, que simulan la percepción visual humana, pero bajo modelos geométricos y computacionales rigurosos [33]. Estos sistemas se componen típicamente de una unidad de adquisición, sensores de imagen y módulos de procesamiento de alto rendimiento. Dependiendo de su configuración, los sistemas de visión pueden operar en modalidad monocular o estereoscópica; esta última permite inferir profundidad mediante triangulación epipolar, a partir de la disparidad observada entre pares de imágenes captadas desde distintas posiciones, lo cual posibilita la estimación de coordenadas tridimensionales de los objetos en la escena [37].

Sus aplicaciones abarcan desde la inspección automatizada en líneas de producción hasta el reconocimiento biométrico y la reconstrucción forense. En entornos urbanos y escenarios reales, donde la detección y el seguimiento de peatones enfrentan desafíos como oclusiones, variaciones de iluminación, lluvia o niebla, los sistemas de visión pasiva basados en cámaras constituyen una alternativa pasiva eficiente, al generar información tridimensional sin necesidad de emisores activos. Aunque los sensores láser presentan mejor desempeño en condiciones de baja iluminación, pueden verse afectados por fenómenos atmosféricos como la niebla o la lluvia intensa [38]. La fusión entre ambos tipos de sensores ha demostrado ser una estrategia eficaz para compensar sus limitaciones individuales, mejorando así la robustez del sistema de percepción, por ejemplo en aplicaciones de seguridad vial.

2.1.2. Sistemas de estereovisión

La estereovisión se fundamenta en la adquisición simultánea de una escena desde dos posiciones ligeramente separadas. A partir de la identificación de correspondencias entre puntos clave en ambas imágenes, se calcula la disparidad y se infiere la profundidad mediante triangulación geométrica. Para garantizar exactitud métrica, es imprescindible realizar una calibración geométrica del sistema estéreo, que permita estimar con precisión tanto los parámetros intrínsecos de cada cámara como la geometría relativa (rotación y traslación) entre ellas [39].

Esta tecnología se ha consolidado como herramienta de percepción 3D pasiva en robótica móvil, reconstrucción estructural y sistemas de asistencia avanzada al conductor (Advanced Driver Assistance Systems, ADAS, por sus siglas en inglés). Frente a tecnologías activas como el LiDAR, la estereovisión representa una alternativa pasiva más económica, con una favorable relación entre resolución espacial y tasa de actualización, particularmente adecuada en entornos con buena textura superficial. Esta técnica permite generar mapas de profundidad mediante disparidad estereoscópica sin depender de emisores activos, lo cual resulta ventajoso en aplicaciones sensibles a interferencias ópticas o restricciones energéticas [40].

La configuración geométrica fundamental de un sistema de estereovisión, incluyendo la disposición de cámaras y la línea base que las separa, se ilustra en la Figura 2.1.

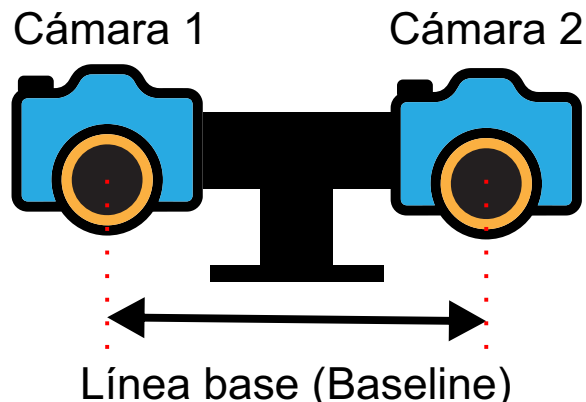


Figura 2.1: Esquema de un sistema de estereovisión y su línea base.

2.1.3. Sistemas láser para visión artificial

Los sensores láser utilizados en visión artificial permiten la reconstrucción tridimensional del entorno mediante distintas técnicas de sensado activo, en las cuales el sistema emite su propia fuente de luz estructurada o modulada. Entre las tecnologías más difundidas se encuentran los sistemas LiDAR (Light Detection and Ranging), que estiman distancias utilizando el principio de tiempo de vuelo (Time-of-Flight, ToF), tal como se ha implementado en sistemas móviles para modelado tridimensional en interiores [35]. Estos pueden operar midiendo el retardo de un pulso láser (ToF pulsado) o el desfase de una señal modulada (ToF por desplazamiento de fase). Por otro lado, están los sistemas basados en interferometría óptica, que emplean la superposición coherente de haces láser para detectar desplazamientos muy pequeños con alta sensibilidad, lo cual ha permitido su aplicación en tareas de inspección superficial, perfilado 3D sin contacto y caracterización estructural de microcomponentes [41].

Por su parte, la triangulación dinámica es una técnica de reconstrucción tridimensional basada en principios geométricos, en la cual se proyecta un haz láser hacia un objeto y se mide el punto reflejado desde una ubicación distinta. En la triangulación dinámica se realiza el posicionamiento angular activo del haz emisor. El ángulo de emisión se controla para dirigir el haz hacia diferentes regiones del espacio, mientras que el ángulo de recepción se determina midiendo el instante angular en que el haz reflejado alcanza el sensor dentro de la apertura giratoria, de modo que ambos ángulos son conocidos en todo momento mediante actuadores y sensores de posición integrados. Conociendo estos ángulos y la distancia base entre emisor y receptor, se puede calcular la ubicación espacial del punto de impacto utilizando la ley de los senos. Esta técnica ofrece alta precisión, siendo adecuada para aplicaciones en visión artificial, inspección industrial y navegación robótica [42].

2.1.4. Sistemas híbridos de visión artificial

Los sistemas híbridos de visión artificial integran sensores ópticos heterogéneos con el objetivo de compensar las limitaciones individuales de cada tecnología. Este enfoque es común en configuraciones que combinan sensores pasivos (como cámaras RGB o estéreo) con sensores activos (como

LiDAR o sistemas de triangulación láser), permitiendo una percepción tridimensional más robusta y precisa. La fusión de información puede realizarse en distintos niveles del pipeline de percepción: a nivel de datos crudos (low-level fusion), a nivel de características (feature-level fusion) o a nivel de decisiones (high-level fusion), dependiendo del grado de procesamiento local y la sincronización temporal entre dispositivos [43].

Aplicaciones como la navegación autónoma, la inspección estructural o la vigilancia inteligente se benefician de esta arquitectura sensorial redundante y robusta. Por ejemplo, la combinación de cámaras RGB con sensores de profundidad activa, como los sistemas basados en luz estructurada o escáneres láser, permite construir mapas tridimensionales texturizados, donde se integran simultáneamente la geometría del entorno y la apariencia visual de la escena. Esta sinergia facilita tareas como la localización, la reconstrucción de entornos y la interacción avanzada con objetos, incluso en tiempo real [44].

En la Figura 2.2 se esquematiza un ejemplo de esta integración, donde se ilustran los respectivos campos de visión —Field of View (FoV)— de dos sensores heterogéneos montados en un sistema móvil.

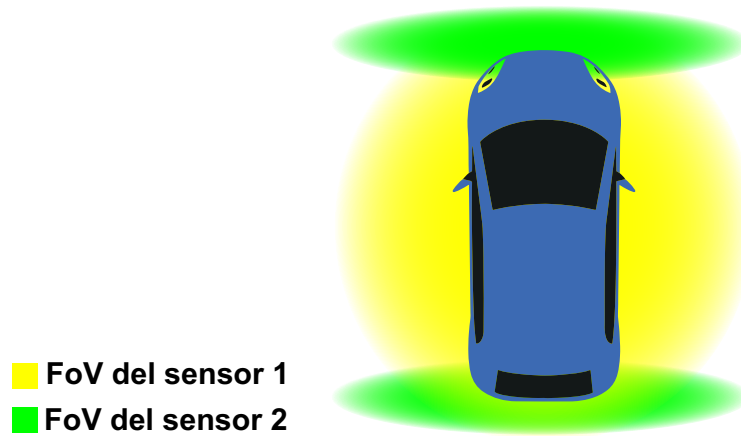


Figura 2.2: Ejemplo esquemático de la superposición de campos de visión (FOV) de dos sensores heterogéneos montados en un sistema móvil.

2.1.5. Sistema híbrido Estéreo-TVS

El sistema desarrollado en este trabajo combina un Sistema de Visión Técnica (TVS), basado en triangulación dinámica, con un sistema de estereovisión compuesto por dos cámaras de alta definición (HD, por sus siglas en inglés). Esta arquitectura híbrida permite integrar la alta precisión geométrica puntual del láser con la reconstrucción tridimensional densa y global de la escena proporcionada por las cámaras estéreo.

Las principales ventajas de esta integración son:

- **Alta precisión local:** El sistema TVS, basado en triangulación dinámica, permite estimaciones tridimensionales puntuales con alta precisión incluso sobre superficies homogéneas, gracias al contraste proporcionado por el haz láser así como la detección angular controlada del punto reflejado [42].
- **Recolección densa de datos visuales:** El sistema de estereovisión permite generar mapas de disparidad sobre grandes áreas mediante algoritmos de correspondencia binocular, lo que facilita reconstrucciones tridimensionales densas. Además, aporta contexto visual a través de imágenes monocromáticas o RGB, dependiendo de la configuración de cámaras utilizada [37].
- **Equilibrio precisión-cobertura:** La fusión de ambos sistemas mejora la densidad de información en zonas críticas, extendiendo la cobertura efectiva de reconstrucción.

2.1.6. Marcos de referencia en visión multisensorial

En sistemas de percepción tridimensional, cada sensor opera dentro de su propio sistema de coordenadas local, también denominado sistema de referencia propio. Para lograr una fusión coherente entre sensores heterogéneos (por ejemplo, estereovisión y sistemas de triangulación láser), es indispensable transformar todas las mediciones al marco común de un sistema de referencia global, o bien trasladar los datos de un sensor al sistema de referencia del otro sensor que se utilizará como base. El manejo riguroso de sistemas de referencia es fundamental para procesos de calibración extrínseca, ya que permite estimar con precisión la relación espacial entre dispositivos de captura

3D montados en distintas configuraciones [45].

Un marco global puede ser un sistema fijo en el entorno, como el asociado a una mesa óptica de laboratorio. Dicho sistema se considera un marco absoluto o inercial, bajo el supuesto de que no está sujeto a aceleraciones significativas. Las posiciones y orientaciones se expresan con respecto a este marco global, mientras que las coordenadas medidas por los sensores se consideran relativas o locales [46].

En la Fig. 2.3 se ilustra un ejemplo de este tipo de configuración, donde se representan tanto el sistema de referencia absoluto como los marcos locales asociados a sensores individuales en un entorno multisensorial.

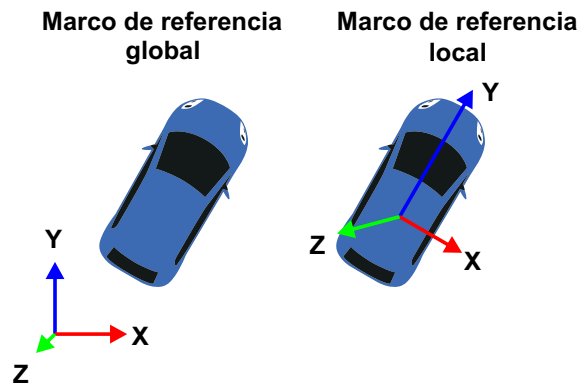


Figura 2.3: Ejemplo de marcos de referencia global y local en un sistema multisensorial.

2.1.7. Calibración extrínseca en sistemas multisensoriales

La calibración extrínseca de un sistema multisensorial consiste en estimar la transformación rígida que relaciona los marcos de referencia de sensores distintos, usualmente expresada como una matriz homogénea $\mathbf{T} \in \mathbb{SE}(3)$. Esta matriz se define como:

$$\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \quad (2.1)$$

La matriz homogénea que define la transformación extrínseca entre dos sistemas de referencia se presenta en la Ecuación (2.1), donde \mathbf{R} denota la rotación y \mathbf{t} la traslación entre los sistemas de

coordenadas. La estimación precisa de esta transformación es fundamental para asegurar la coherencia geométrica en tareas de percepción tridimensional y fusión sensorial. Errores en los parámetros extrínsecos pueden generar desalineamientos sistemáticos entre datos capturados por sensores complementarios, comprometiendo la calidad de la reconstrucción de la escena y la fiabilidad de procesos posteriores como la detección de objetos o la navegación autónoma [47].

2.2. Sistema de Visión Técnica (TVS)

El Sistema de Visión Técnica (TVS) es un dispositivo diseñado para la medición tridimensional de superficies mediante el método de Triangulación Dinámica [12]. Este sistema se compone de dos subsistemas principales: el posicionador y la apertura, los cuales deben estar alineados con precisión, ya que su configuración determina la exactitud del proceso de triangulación láser.

La técnica de triangulación se basa en la proyección de un haz láser desde el posicionador hacia la superficie de interés. Cuando el haz incide sobre la superficie, se genera un punto de reflexión cuya luz es capturada por el subsistema de apertura. A partir de los ángulos de emisión y recepción, junto con la separación entre ambos subsistemas, se pueden calcular las coordenadas bidimensionales del punto de impacto del láser en la superficie con base en principios de trigonometría. En la figura 2.4 se ilustra la disposición de los componentes del sistema TVS y los componentes de la triangulación. En las siguientes subsecciones se detallan los componentes que conforman ambos subsistemas, el posicionador y la apertura.

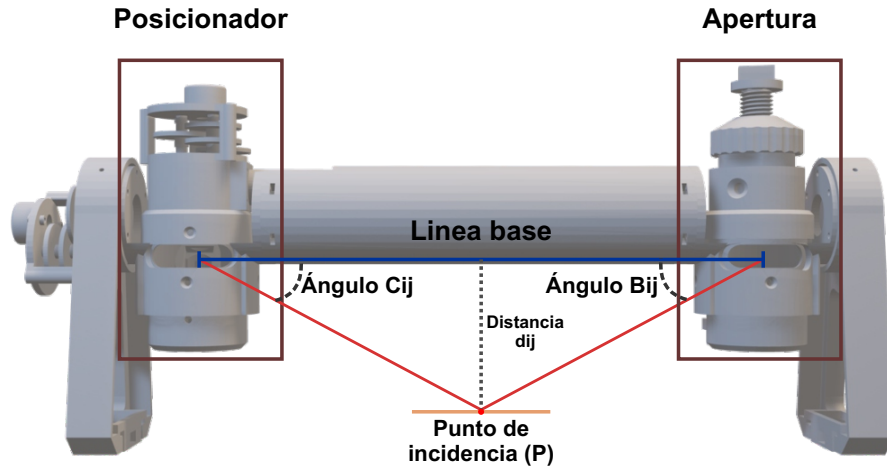


Figura 2.4: Esquema general del Sistema de Visión Técnica (TVS)

2.2.1. Subsistema de posicionador

El posicionador es el componente responsable de dirigir el haz de luz hacia los puntos de interés en la superficie a escanear. Para ello, emplea varios elementos clave que garantizan un control preciso sobre la trayectoria del haz.

El sistema posicionador opera reflejando el haz láser en dos espejos inclinados a 45° , los cuales redirigen la luz hacia la salida. La orientación de estos espejos es controlada mediante un motor a pasos, cuya posición angular es determinada con precisión al estar caracterizada la relación entre el número de pasos y el desplazamiento angular resultante. Para mejorar la resolución del sistema, el motor a pasos está acoplado a una transmisión de engranajes.

En la figura 2.5 se muestra el subsistema posicionador junto con sus principales componentes.

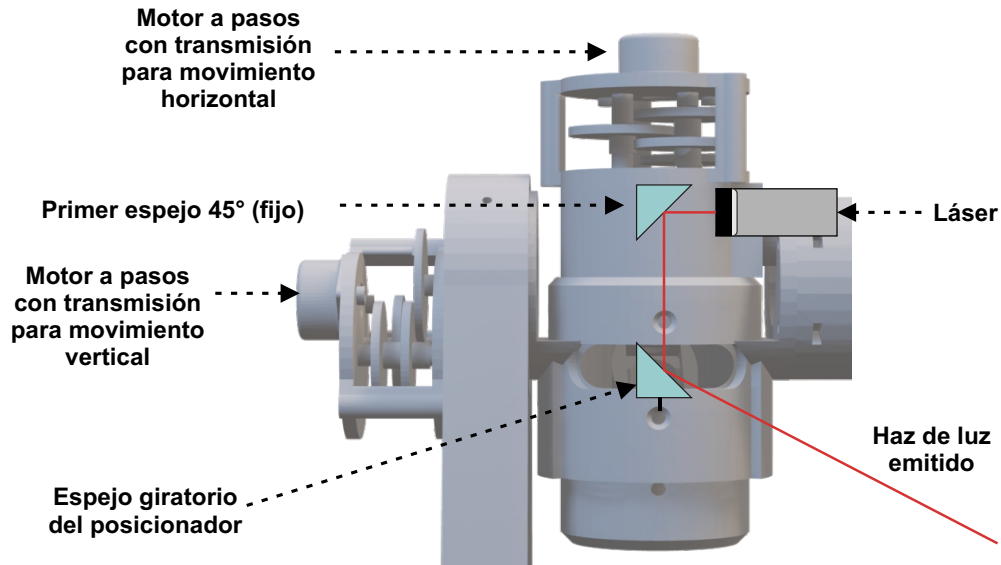


Figura 2.5: Componentes del subsistema posicionador del TVS

2.2.2. Subsistema de apertura

El subsistema de apertura del TVS es el encargado de recibir el haz de luz reflejado desde la superficie medida. El sistema cuenta con un espejo rotatorio inclinado a 45°, el cual captura el haz reflejado y lo redirige hacia un fotodiodo que actúa como sensor. El fotodiodo convierte la luz en una señal eléctrica, que es enviada al microcontrolador encargado del procesamiento de la señal.

El espejo rotatorio es accionado por un motor de corriente continua (DC), el cual está acoplado a un disco con un orificio. Este orificio es detectado mediante un optointerruptor, lo que permite determinar la posición angular del espejo. El orificio marca el inicio de cada revolución del espejo, es decir, el ángulo de referencia de 0°. Al medir la diferencia entre esta referencia y el instante en que el fotodiodo detecta el haz de luz, es posible calcular la posición angular del espejo de apertura en el momento de la detección.

En la figura 2.6 se muestra el subsistema de apertura del TVS junto con sus principales componentes.

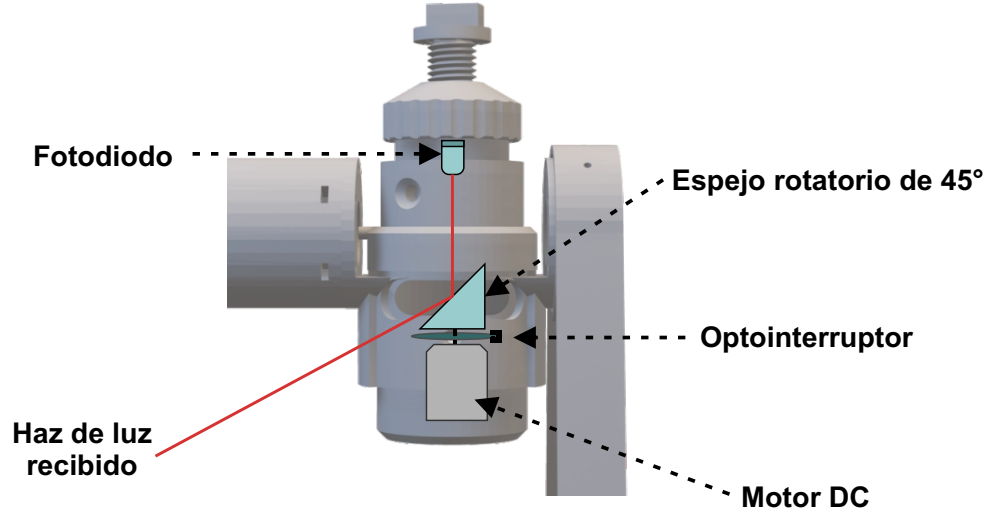


Figura 2.6: Componentes del subsistema apertura del TVS

2.2.3. Triangulación dinámica y nube de puntos 3D

Conociendo el ángulo de inclinación del espejo del posicionador, es posible determinar completamente las coordenadas tridimensionales del punto de incidencia. A partir de la ley de senos, la componente en el eje x se obtiene como se muestra en la Ecuación (2.2):

$$x_{ij} = a \frac{\sin B_{ij} \cdot \sin C_{ij} \cdot \cos \sum_{j=1}^j \beta_j}{\sin[180^\circ - (B_{ij} + C_{ij})]} \quad (2.2)$$

La coordenada en y se calcula considerando el valor del ángulo B_{ij} , y está definida por tramos como se observa en la Ecuación (2.3):

$$y_{ij} = \begin{cases} a \left(\frac{1}{2} - \frac{\sin B_{ij} \cdot \cos C_{ij}}{\sin[180^\circ - (B_{ij} + C_{ij})]} \right), & \text{si } B_{ij} \leq 90^\circ, \\ -a \left(\frac{1}{2} + \frac{\sin B_{ij} \cdot \cos C_{ij}}{\sin[180^\circ - (B_{ij} + C_{ij})]} \right), & \text{si } B_{ij} \geq 90^\circ. \end{cases} \quad (2.3)$$

Finalmente, la coordenada en z se obtiene mediante la expresión mostrada en la Ecuación (2.4):

$$z_{ij} = a \frac{\sin B_{ij} \cdot \sin C_{ij} \cdot \sin \sum_{j=1}^j \beta_j}{\sin[180^\circ - (B_{ij} + C_{ij})]} \quad (2.4)$$

El método de triangulación dinámica permite realizar un barrido tanto en dirección horizontal como vertical mediante los motores a pasos que controlan la orientación del posicionador. Mientras tanto, el espejo del subsistema de apertura rota continuamente, permitiendo la captura de múltiples puntos en la escena. Este proceso genera una nube de puntos tridimensional que representa la superficie escaneada.

2.3. Calibración del sistema TVS

En términos generales, la calibración de un instrumento de medición se define como el conjunto de operaciones que establecen, bajo condiciones específicas, la relación entre los valores indicados por el instrumento y los valores correspondientes de una magnitud medida mediante un patrón de referencia trazable [48]. Esta definición, de raíz metrológica, enfatiza la necesidad de una comparación directa con estándares certificados para validar la exactitud del sistema.

En la práctica, no siempre es posible contar con patrones físicos trazables para llevar a cabo una calibración formal bajo estándares metrológicos. Esto ocurre, por ejemplo, en sistemas embebidos, entornos de difícil acceso o configuraciones experimentales que no permiten el uso de artefactos de referencia certificados. En tales casos, se recurre a enfoques alternativos que permiten caracterizar el comportamiento geométrico del sistema a partir de condiciones estructurales del entorno o mediante la comparación con un segundo sistema auxiliar de medición. En el presente trabajo, se emplea un sistema de visión estéreo como referencia tridimensional externa, lo que permite estimar los parámetros geométricos del TVS de forma indirecta, apoyándose en la observación conjunta de superficies planas y trayectorias proyectadas [21].

Este enfoque es coherente con la definición adoptada en visión artificial, donde la calibración de una cámara implica la estimación de los parámetros geométricos que rigen la relación entre el espacio tridimensional y su representación en imágenes. En particular, se distingue entre:

- **Parámetros intrínsecos:** Propiedades internas del sistema, tales como distancias focales, ejes ópticos, alineaciones internas y factores de escala específicos del dispositivo.

- **Parámetros extrínsecos:** Transformaciones geométricas que relacionan el sistema con marcos de referencia externos, generalmente representadas por una matriz de rotación $\mathbf{R} \in SO(3)$ y un vector de traslación $\mathbf{t} \in \mathbb{R}^3$ [49]. Estos parámetros definen cómo se posiciona y orienta el sensor con respecto a un marco global o a otro sensor, y son esenciales en arquitecturas multisensoriales para garantizar la coherencia espacial entre las distintas fuentes de datos.

En el caso del Sistema de Visión Técnica (TVS), los parámetros intrínsecos no se limitan a elementos ópticos, sino que incluyen características mecánicas y geométricas que afectan directamente el proceso de triangulación dinámica, tales como:

- La inclinación individual y la alineación entre los ejes del espejo del posicionador y el espejo en la apertura.
- La sincronización angular entre el punto de activación del optointerruptor y la orientación del espejo en la apertura.
- La relación entre los pasos del motor y el desplazamiento efectivo del haz láser.

Dado que estos factores inciden en la determinación de coordenadas 3D, su calibración es indispensable para garantizar la fidelidad geométrica del sistema.

En la siguiente sección se presentará la calibración de cámaras convencionales, abordando tanto la estimación de parámetros intrínsecos como los modelos de distorsión óptica. Posteriormente, se retomará el concepto de calibración extrínseca, la cual es fundamental en entornos multisensoriales, ya que permite establecer transformaciones precisas entre el Sistema de Visión Técnica (TVS) y el sistema de estereovisión.

2.4. Calibración de la Cámara

Considerando que el sistema híbrido propuesto incorpora un subsistema de estereovisión, resulta indispensable comprender los fundamentos teóricos y geométricos asociados a la calibración de cámaras, tanto de forma individual como en configuración estéreo. Aunque este trabajo no se

centra en el desarrollo de un nuevo método de calibración para cámaras, esta sección introduce los conceptos fundamentales necesarios para estimar los parámetros intrínsecos y extrínsecos que modelan el proceso de proyección. Dichos parámetros son imprescindibles para emplear el sistema estéreo como herramienta auxiliar en la calibración del sistema híbrido. Además, varios de estos principios pueden extenderse, mediante analogías, al proceso de calibración del TVS y del sistema combinado.

En visión por computadora, la calibración de una cámara es el proceso de estimar sus parámetros intrínsecos y extrínsecos para modelar la transformación geométrica y óptica que relaciona las coordenadas tridimensionales del mundo con sus proyecciones en un sensor bidimensional [14]. Este procedimiento es fundamental en aplicaciones como la reconstrucción 3D, la metrología y la navegación autónoma.

2.4.1. Modelos de cámara

El modelado de una cámara es esencial en visión por computadora, ya que permite describir matemáticamente la proyección de un entorno tridimensional en un sensor bidimensional [50]. Existen diversos modelos ópticos que representan este proceso, cada uno con diferentes niveles de aproximación a la realidad, desde el modelo pinhole ideal hasta formulaciones que incorporan parámetros de distorsión, orientación espacial y desplazamiento del centro óptico.

Modelo Pinhole

Un modelo de cámara ampliamente utilizado es el modelo de cámara pinhole, el cual idealiza la formación de imágenes mediante un pequeño orificio sin lentes. En este modelo, la proyección de un punto tridimensional en la imagen se describe mediante una transformación proyectiva utilizando una matriz de calibración [51]. Debido a su simplicidad matemática y la ausencia de distorsión óptica, este modelo es la base para la mayoría de los algoritmos de calibración en visión artificial y reconstrucción 3D.

En el modelo pinhole, los rayos de luz atraviesan un pequeño orificio sin lentes y se proyectan

en un plano de imagen invertido. Este modelo idealiza la proyección sin aberraciones ópticas ni distorsiones geométricas. La figura 2.7 ilustra este principio.

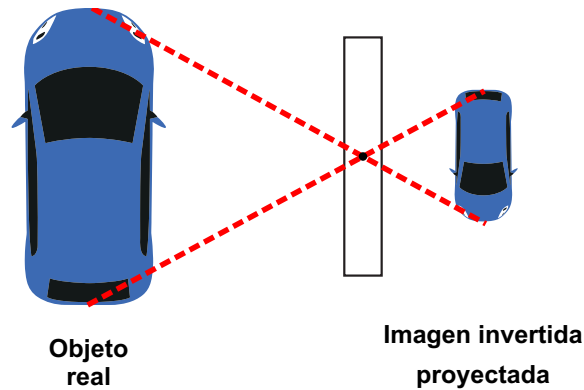


Figura 2.7: Esquema del modelo de cámara pinhole y su proyección invertida.

Sin embargo, las cámaras reales incluyen lentes que introducen distorsiones ópticas que afectan la precisión geométrica de las imágenes. Para modelar estas aberraciones, se utilizan funciones de distorsión radial y tangencial, las cuales representan desplazamientos sistemáticos respecto a la posición ideal de los puntos proyectados. Estos efectos deben ser corregidos mediante algoritmos de calibración que integran modelos como el de Brown-Conrady, ampliamente utilizado por su capacidad para representar con precisión las deformaciones inducidas por el sistema óptico [52].

2.4.2. Transformación de coordenadas en la imagen

El proceso de proyección de un punto tridimensional a una imagen se descompone en dos transformaciones principales:

1. **Transformación del mundo a la cámara:** Se describe mediante los parámetros extrínsecos de la cámara, que incluyen una matriz de rotación R y un vector de traslación \mathbf{t} [53].
2. **Proyección de la cámara al plano de imagen:** Se realiza mediante la matriz intrínseca M_{int} , que encapsula los parámetros ópticos de la cámara [14].

La Figura 2.8 muestra la geometría de la proyección de un punto del mundo en la imagen, ilustrando la relación entre los marcos de coordenadas del mundo y de la cámara, así como las principales

variables involucradas: las coordenadas del punto P en el sistema del mundo $P_m = (X_m, Y_m, Z_m)$, sus coordenadas en el sistema de la cámara $P_c = (X_c, Y_c, Z_c)$, el punto proyectado en el plano de imagen $P_i = (X_i, Y_i)$, la distancia focal f , y el centro óptico de la cámara C_m expresado en coordenadas del mundo.

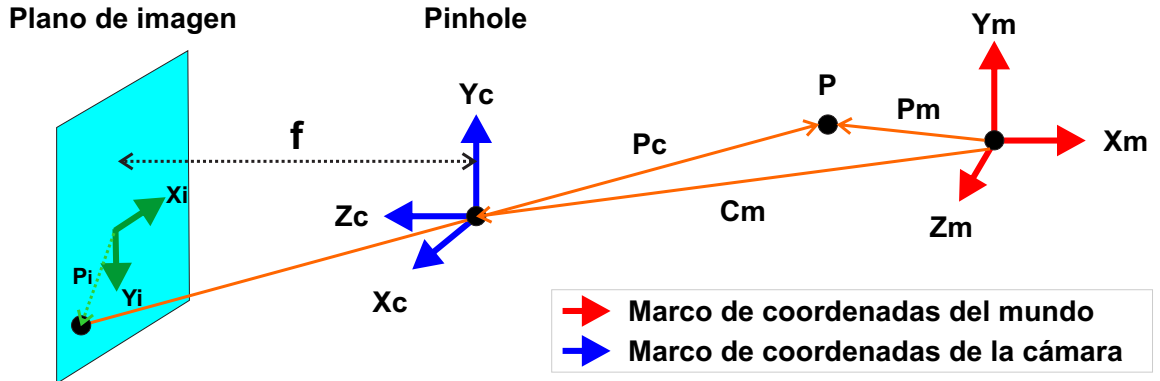


Figura 2.8: Modelo geométrico de la proyección de un punto en el mundo a coordenadas de la cámara.

2.4.3. Matriz extrínseca: transformación del mundo a la cámara

En visión por computadora, la matriz extrínseca M_{ext} se utiliza para transformar las coordenadas de un punto en el sistema de referencia del mundo al sistema de coordenadas de la cámara [54].

Esta matriz está compuesta por dos elementos fundamentales:

- R , una matriz de rotación 3×3 que describe la orientación de la cámara con respecto al sistema de referencia del mundo.
- t , un vector de traslación 3×1 que define la posición de la cámara en el espacio del mundo.

Matemáticamente, la transformación entre un punto en el espacio del mundo P_m y su representación en el sistema de la cámara P_c se expresa mediante una transformación rígida, como se muestra en la Ecuación (2.5):

$$P_c = RP_m + t \quad (2.5)$$

Expandiendo esta ecuación en forma matricial, se obtiene la Ecuación (2.6):

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (2.6)$$

Para simplificar la aplicación de esta transformación, se emplea una representación en coordenadas homogéneas, lo que permite expresar la operación mediante una única multiplicación matricial. En este caso, la matriz extrínseca adopta la forma de la Ecuación (2.7):

$$M_{\text{ext}} = \begin{bmatrix} R & t \\ 0_{1 \times 3} & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.7)$$

Dado que R es una matriz ortonormal, cumple:

$$R^T R = I, \quad R^{-1} = R^T,$$

lo cual permite invertir transformaciones extrínsecas de forma directa. En particular, si se desea expresar un punto del sistema de coordenadas de la cámara al sistema del mundo, la transformación inversa se expresa como:

$$P_m = R^T (P_c - t),$$

lo que evita realizar una inversión matricial explícita y garantiza la consistencia geométrica entre marcos de referencia.

2.4.4. Sensor de la cámara y conversión de coordenadas

En la práctica, la conversión de coordenadas métricas del plano de imagen a coordenadas en píxeles se encapsula en la matriz intrínseca, la cual integra la distancia focal (en unidades de píxel) y la ubicación del punto principal. Esta matriz permite traducir la geometría óptica a una representación discreta acorde con la resolución del sensor, sin necesidad de modelar explícitamente el tamaño físico del sensor ni las dimensiones de cada píxel.

2.4.5. Matriz intrínseca: conversión a píxeles

La matriz intrínseca de una cámara es un modelo matemático que encapsula los parámetros internos del sistema óptico, tales como la distancia focal y el punto principal, permitiendo convertir coordenadas métricas del plano de imagen a coordenadas en píxeles dentro del sensor [55]. Esta transformación es fundamental en la calibración de cámaras, ya que habilita el uso de las imágenes en tareas de procesamiento digital, como la reconstrucción tridimensional o la percepción artificial. Esta matriz se define como se muestra en la Ecuación (2.8):

$$M_{\text{int}} = \begin{bmatrix} f_x & s & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.8)$$

donde:

- f_x y f_y son las distancias focales expresadas en unidades de píxeles. Se relacionan con la distancia focal real f y el tamaño del píxel en cada eje mediante la Ecuación (2.9):

$$f_x = \frac{f}{s_x}, \quad f_y = \frac{f}{s_y} \quad (2.9)$$

donde s_x y s_y representan los tamaños del píxel en las direcciones x e y , respectivamente.

- s es el factor de sesgo (skew), que modela la inclinación entre los ejes del sensor. En la mayoría de los sistemas de visión modernos, este valor es cero.
- o_x y o_y representan las coordenadas del punto principal, es decir, la proyección del centro óptico en el sensor de la cámara.

2.4.6. Matriz de proyección

La matriz de proyección P_{proj} se obtiene como el producto entre la matriz intrínseca M_{int} y la matriz extrínseca M_{ext} , como se observa en la Ecuación (2.10):

$$P_{proj} = M_{int}M_{ext} \quad (2.10)$$

Al expandir los términos, la matriz toma la siguiente forma explícita, mostrada en la Ecuación (2.11):

$$P_{proj} = \begin{bmatrix} f_x r_{11} + s r_{21} & f_x r_{12} + s r_{22} & f_x r_{13} + s r_{23} & f_x t_x + s t_y + o_x \\ f_y r_{21} & f_y r_{22} & f_y r_{23} & f_y t_y + o_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \quad (2.11)$$

2.4.7. Proceso de calibración experimental

El proceso de calibración de la cámara consiste en estimar los parámetros de la matriz de proyección mediante un conjunto de correspondencias entre puntos en el mundo real y su proyección en la imagen [15]. Para ello, se parte de un objeto de geometría conocida, como un tablero de ajedrez o un cubo con un patrón visual, que permite identificar puntos con coordenadas tridimensionales $\mathbf{P}_m = (X_m, Y_m, Z_m)$ en un sistema de referencia del mundo.

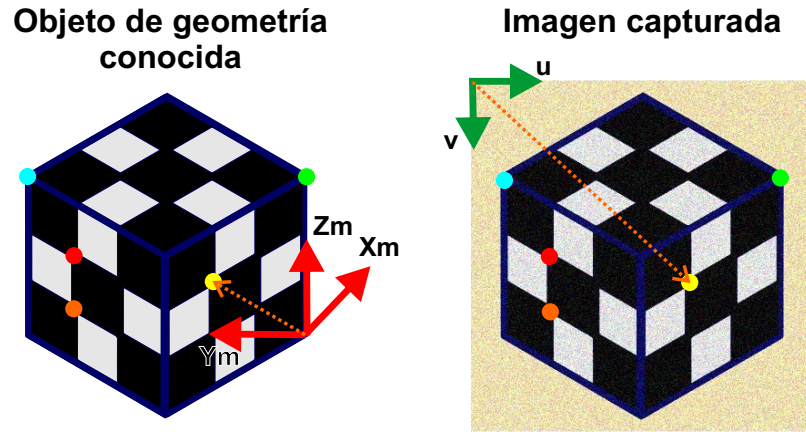


Figura 2.9: Ejemplo de objeto de calibración con un patrón visual. Este patrón permite establecer correspondencias entre puntos 3D en el mundo real y su proyección en la imagen.

2.5. Métodos de estimación de la matriz de proyección

La estimación precisa de la matriz de proyección $P_{proj} = M_{int}M_{ext}$ es esencial para aplicaciones de reconstrucción 3D, percepción multisensorial y sistemas de visión estéreo. Existen múltiples enfoques que permiten inferir los parámetros intrínsecos y extrínsecos de una cámara, cada uno con diferentes grados de complejidad computacional y precisión geométrica. Asimismo, las distorsiones ópticas introducidas por las lentes deben ser modeladas y compensadas antes de cualquier operación de triangulación o correspondencia estéreo. A continuación se describen los principales métodos y modelos relevantes.

2.5.1. Método DLT (Direct Linear Transform)

El método DLT es un enfoque algebraico clásico que permite estimar la matriz de proyección a partir de correspondencias entre puntos en el mundo tridimensional P_m y sus proyecciones en la imagen P_i . Se construye un sistema homogéneo $A\mathbf{p} = 0$, donde \mathbf{p} es el vector de parámetros de la matriz P_{proj} , y la solución se obtiene mediante descomposición en valores singulares (SVD) [56].

Este método es lineal, no requiere optimización iterativa y es adecuado para una primera estimación. Sin embargo, presenta limitaciones significativas: no modela distorsiones ópticas, es sensible al ruido en las correspondencias, y la calidad de la estimación se ve fuertemente afectada por la configuración

espacial de los puntos (por ejemplo, si son coplanares).

2.5.2. Método de Zhang

El método propuesto en [15] es ampliamente utilizado debido a su precisión y aplicabilidad práctica. Permite la calibración intrínseca y extrínseca de una cámara a partir de varias vistas de un patrón plano (como un tablero de ajedrez) colocado en diferentes orientaciones. Se basa en dos etapas:

1. Estimación inicial de homografías 2D a partir de cada vista del plano, que permiten derivar los parámetros intrínsecos de la cámara.
2. Optimización conjunta de todos los parámetros mediante ajuste por mínimos cuadrados del error de reproyección (bundle adjustment), usualmente implementado con el algoritmo de Levenberg-Marquardt.

Este método también estima coeficientes de distorsión radial y tangencial, y está implementado en bibliotecas como OpenCV y MATLAB.

2.6. Modelado de distorsión óptica

Las cámaras reales presentan distorsiones ópticas que deben modelarse para obtener proyecciones geoméricamente precisas. Estas distorsiones afectan la precisión en tareas como la correspondencia estéreo, la estimación de profundidad y la calibración multisensorial.

2.6.1. Rectificación de imágenes

En sistemas de visión artificial, las aberraciones introducidas por los componentes ópticos deben ser corregidas para garantizar una representación geoméricamente precisa del entorno. La rectificación óptica compensa estas distorsiones, generando imágenes equivalentes a las capturadas por una cámara ideal basada en el modelo pinhole.

Este proceso es esencial en tareas como la visión estéreo, la reconstrucción 3D y la metrología computacional. En particular, la rectificación garantiza que líneas rectas del mundo real no se

proyecten como curvas en la imagen digital.



Figura 2.10: Ejemplo del efecto de la distorsión radial sobre una imagen y su corrección mediante rectificación.

2.7. Calibración de un sistema de estereovisión

La visión estereoscópica es una técnica ampliamente utilizada en el campo de la visión por computadora para inferir la estructura tridimensional de una escena a partir de dos o más imágenes capturadas desde distintas posiciones [57].

Este método se basa en el principio de triangulación estéreo, el cual permite calcular la profundidad de un punto en el espacio al establecer su correspondencia en dos vistas diferentes. Dicha correspondencia se aprovecha para estimar la componente de profundidad, reforzada por la geometría epipolar [50], que restringe el espacio de búsqueda de coincidencias.

En un sistema de estereovisión, se dispone de dos cámaras que se encuentran separadas por una distancia conocida b , a la cual se le denomina "línea base"(baseline). Las proyecciones de un punto tridimensional (x, y, z) en los planos de imagen de ambas cámaras pueden describirse a través de las siguientes expresiones, asumiendo una configuración canónica (es decir, con la cámara derecha trasladada respecto a la izquierda a lo largo del eje X y sin rotación relativa) [58]:

$$u_l = f_x \frac{x}{z} + o_x, \quad v_l = f_y \frac{y}{z} + o_y \quad (2.12)$$

$$u_r = f_x \frac{x - b}{z} + o_x, \quad v_r = f_y \frac{y}{z} + o_y \quad (2.13)$$

Dado que estas proyecciones se refieren al mismo punto en el espacio, se define la disparidad como la diferencia en la coordenada horizontal entre ambas imágenes [59], como se muestra en la Ecuación (2.14):

$$d = u_l - u_r \quad (2.14)$$

La profundidad z se deduce mediante la siguiente ecuación de triangulación, mostrada en la Ecuación (2.15):

$$z = \frac{bf_x}{d} \quad (2.15)$$

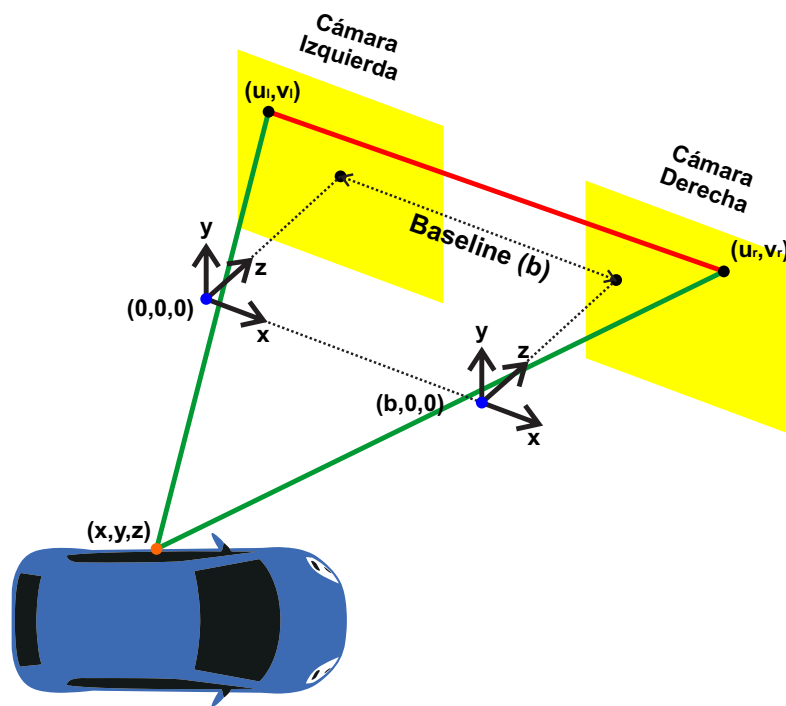


Figura 2.11: Triangulación con un sistema estereoscópico.

No obstante, en la práctica, las cámaras no siempre se ubican en configuraciones perfectas ni copla-
nares. Por ello, la calibración estéreo permite determinar con precisión los parámetros extrínsecos
(posición y orientación relativa entre cámaras) y, en caso necesario, refinar los parámetros intrínsecos
de cada cámara (distancia focal, centro óptico y coeficientes de distorsión) previamente obtenidos

por calibración individual [60]. Este proceso suele implicar el uso de algoritmos de optimización no lineal, como *bundle adjustment*, cuyo objetivo es minimizar el *error de reproyección* entre los puntos de imagen observados y los proyectados por el modelo de cámara. Una estrategia ampliamente utilizada es el algoritmo de Levenberg–Marquardt, que permite resolver este problema iterativo de mínimos cuadrados con robustez numérica [61]. Una vez completada esta etapa de refinamiento, las mediciones de profundidad obtenidas por triangulación permiten reconstrucciones tridimensionales precisas de la escena.

Para lograr esto, es esencial identificar los puntos homólogos o correspondientes, es decir, aquellos píxeles en ambas imágenes que se originan en la misma localización física del espacio [62]. La geometría epipolar describe la forma en que un punto en una imagen define una línea epipolar en la otra, reduciendo drásticamente la región de búsqueda de correspondencias y minimizando las ambigüedades [50]. Una vez que se conoce la disparidad d de cada punto (Ecuación (2.14)), se puede estimar su profundidad utilizando la Ecuación (2.15) y reconstruir así la escena en tres dimensiones.

2.7.1. Geometría epipolar

La geometría epipolar es el estudio de la relación geométrica entre las proyecciones de la misma escena en dos imágenes obtenidas desde diferentes puntos de vista [50]. Se utiliza en visión estereoscópica, ya que permite establecer restricciones precisas en la búsqueda de correspondencias entre imágenes, facilitando tanto la calibración de las cámaras como la reconstrucción tridimensional.

En sistemas de visión estereoscópica, durante el proceso de calibración se determina la relación espacial entre las cámaras. La geometría epipolar se utiliza para limitar la búsqueda de la correspondencia de un punto en la imagen izquierda a una línea específica en la imagen derecha, conocida como línea epipolar. Esta restricción reduce significativamente el área de búsqueda, mejorando la robustez y eficiencia en el emparejamiento de puntos.

Matemáticamente, si se tiene un punto u_l en la imagen izquierda, su punto correspondiente u_r en la imagen derecha debe satisfacer la condición epipolar, expresada en la Ecuación (2.16):

$$u_r^T F u_l = 0 \quad (2.16)$$

donde F es la matriz fundamental, que encapsula la relación geométrica entre ambas imágenes sin requerir información de calibración interna de las cámaras.

La Figura 2.12 ilustra esta relación, mostrando cómo un punto en la imagen izquierda, u_l , determina una línea epipolar en la imagen derecha, sobre la cual debe encontrarse el punto correspondiente u_r .

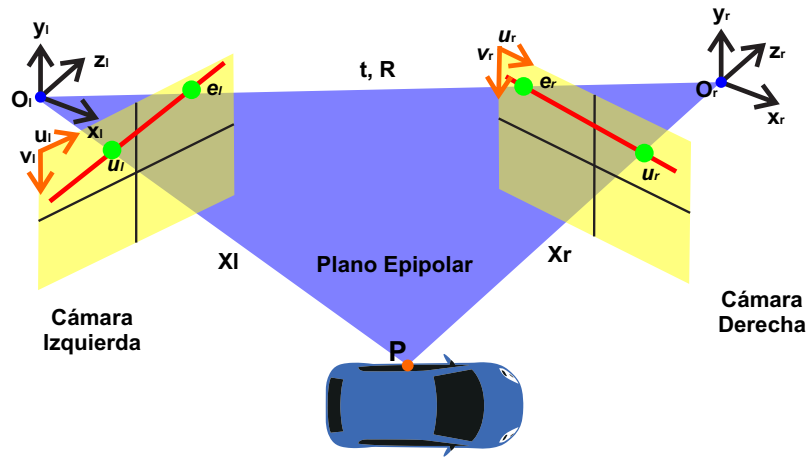


Figura 2.12: Geometría epipolar: relación entre puntos de imagen y líneas epipolares en visión estéreo.

2.7.2. Correspondencia estereoscópica

El problema de la correspondencia estereo radica en determinar qué píxeles (u_l, v_l) en la imagen izquierda se corresponden con qué píxeles (u_r, v_r) en la imagen derecha [63]. Tras un proceso de rectificación estereo, la búsqueda de puntos homólogos se simplifica a la exploración sobre líneas epipolares, reduciendo el problema a la imposición de que las coordenadas verticales sean equivalentes en ambas imágenes, como se indica en la Ecuación (2.17):

$$v_l = v_r \quad (2.17)$$

Los métodos clásicos de correspondencia estereo suelen emplear métricas fotométricas como la suma

de diferencias absolutas (SAD, por sus siglas en inglés), la suma de diferencias al cuadrado (SSD, por sus siglas en inglés) o la correlación cruzada normalizada (NCC, por sus siglas en inglés) para evaluar la similitud entre parches en mapas de disparidad densos [64]. Por otro lado, en el presente trabajo, para la calibración del sistema híbrido no se recurre a este tipo de enfoques. En su lugar, se emplean correspondencias punto a punto obtenidas automáticamente durante la captura de datos. Cada que el TVS toma una lectura, se hace una captura con las cámaras de la mancha del láser, y se segmenta en las imágenes de ambas vistas, para posteriormente calcular su centroide y hacer triangulación estéreo, lo cual permite una correspondencia directa y eficiente entre ambos sistemas. La Figura 2.13 ejemplifica el principio de correspondencia estéreo en imágenes rectificadas, donde un punto observado en la imagen izquierda se busca sobre la misma línea horizontal en la imagen derecha, siguiendo la restricción epipolar.

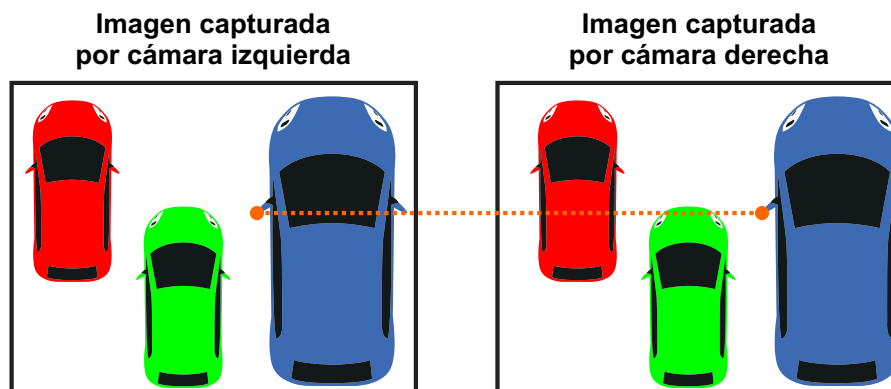


Figura 2.13: Búsqueda de correspondencias en un sistema estéreo rectificado.

2.7.3. Matriz fundamental

La matriz fundamental F es una matriz 3×3 que codifica la relación geométrica entre dos imágenes de una escena [49]. Su obtención se basa en correspondencias de puntos entre las imágenes y en la geometría epipolar. Se representa generalmente como:

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}. \quad (2.18)$$

La condición epipolar se expresa de forma algebraica para cualquier par de puntos homólogos u_l y u_r (en coordenadas homogéneas), como se mostró en la Ecuación (2.16):

Esta ecuación implica que, dada la correspondencia u_l en la imagen izquierda, el punto u_r en la imagen derecha debe ubicarse en la línea epipolar $l_r = Fx_l$, lo que reduce considerablemente la región de búsqueda durante la tarea de emparejamiento.

2.7.4. Matriz esencial

La matriz esencial E codifica la relación geométrica entre dos imágenes en un sistema calibrado, incorporando los parámetros intrínsecos de cada cámara. Se obtiene a partir de la matriz fundamental F y de las matrices de parámetros intrínsecos individuales de la cámara izquierda K_l y derecha K_r [49]. Matemáticamente, se expresa como se indica en la Ecuación (2.19):

$$E = K_r^T F K_l \quad (2.19)$$

donde las matrices intrínsecas se representan generalmente como en la Ecuación (2.20):

$$K_l = \begin{bmatrix} f_{x_l} & 0 & o_{x_l} \\ 0 & f_{y_l} & o_{y_l} \\ 0 & 0 & 1 \end{bmatrix}, \quad K_r = \begin{bmatrix} f_{x_r} & 0 & o_{x_r} \\ 0 & f_{y_r} & o_{y_r} \\ 0 & 0 & 1 \end{bmatrix} \quad (2.20)$$

La matriz esencial E , definida en la Ecuación (2.19), transforma la relación epipolar expresada por F al espacio euclidiano, eliminando las distorsiones introducidas por los parámetros intrínsecos. Esto es fundamental para lograr una reconstrucción tridimensional precisa en entornos calibrados.

2.7.5. Descomposición de la matriz esencial (rotación y traslación)

La matriz esencial E codifica la geometría epipolar entre dos vistas en un sistema estéreo calibrado, encapsulando la rotación R y la traslación t relativas entre ambas cámaras. A partir de E , es posible recuperar R y t —definidos hasta un factor de escala— mediante una descomposición en valores singulares (SVD), lo cual permite reconstruir la transformación rígida entre los marcos de referencia de las cámaras [65].

Este procedimiento genera cuatro posibles combinaciones de rotación y traslación. Para resolver esta ambigüedad, se aplica la *condición de cheirality*, la cual establece que los puntos triangulados deben tener coordenadas de profundidad positivas en ambas cámaras. Esta verificación asegura que la reconstrucción tridimensional obtenida sea físicamente coherente con la configuración real del sistema estéreo.

2.7.6. Calibración extrínseca de dos sistemas

La calibración extrínseca es el proceso mediante el cual se determina la transformación rígida (rotación y traslación) que permite expresar las mediciones de dos sensores distintos en un mismo sistema de coordenadas. En sistemas de percepción multivisual, como aquellos que combinan cámaras con sensores activos de profundidad, dicha transformación es fundamental para una fusión espacial coherente de los datos adquiridos. Específicamente, en configuraciones híbridas como las que integran un sistema de visión estéreo con un sistema de visión técnica (TVS), la estimación precisa de esta transformación permite asociar puntos 3D generados por diferentes subsistemas y mejorar la consistencia métrica de las mediciones.

Los métodos de calibración extrínseca se clasifican habitualmente en función del tipo de escena utilizada y del grado de intervención humana requerido. En cuanto a la escena, se distinguen los métodos *target-based*, que emplean patrones físicos artificiales fácilmente detectables por ambos sensores (como tableros de ajedrez o marcadores AprilTag), y los métodos *target-less*, que aprovechan estructuras presentes naturalmente en el entorno (como paredes, esquinas o planos). En cuanto a la

intervención, existen métodos *manuales*, donde el operador debe seleccionar o verificar correspondencias visuales entre sensores, y métodos *automáticos*, que emplean algoritmos de procesamiento de datos para extraer correspondencias sin intervención externa [21].

Una tercera dimensión importante para clasificar los métodos de calibración extrínseca radica en el tipo de restricción geométrica utilizada para alinear la información entre sensores. Estas restricciones pueden adoptar múltiples formas: punto-a-punto, punto-a-plano, punto-a-línea, línea-a-plano, plano-a-plano, e incluso alineaciones más complejas basadas en movimiento o en medidas de información mutua. Dentro de esta clasificación, los métodos basados en restricciones *plano-a-plano* son particularmente robustos frente a ruido y errores locales, ya que operan sobre entidades geométricas extendidas y permiten inferir la rotación mediante análisis de orientación de superficies [24], [26], [27].

En este trabajo, se propone un método que puede ubicarse dentro de la categoría de calibración *target-less automática* con restricción *plano-a-plano*, según la clasificación planteada en la literatura [21]. A diferencia de los enfoques basados en patrones artificiales, se explota la visibilidad compartida de una superficie estructural plana, sin necesidad de marcas físicas ni texturas distintivas. Las correspondencias espaciales se generan automáticamente a partir de detecciones consistentes del mismo punto físico observadas por ambos sistemas, lo cual permite aplicar modelos geométricos robustos para estimar la transformación extrínseca entre el TVS y el sistema estéreo.

La Figura 2.14 presenta una clasificación jerárquica de los métodos de calibración extrínseca entre sensores, diferenciados por el uso de patrones artificiales y el nivel de automatización. La propuesta de este trabajo, basada en restricciones geométricas naturales sin patrones explícitos, se ubica en la categoría *target-less automático*, la cual ha sido destacada mediante un color verde.

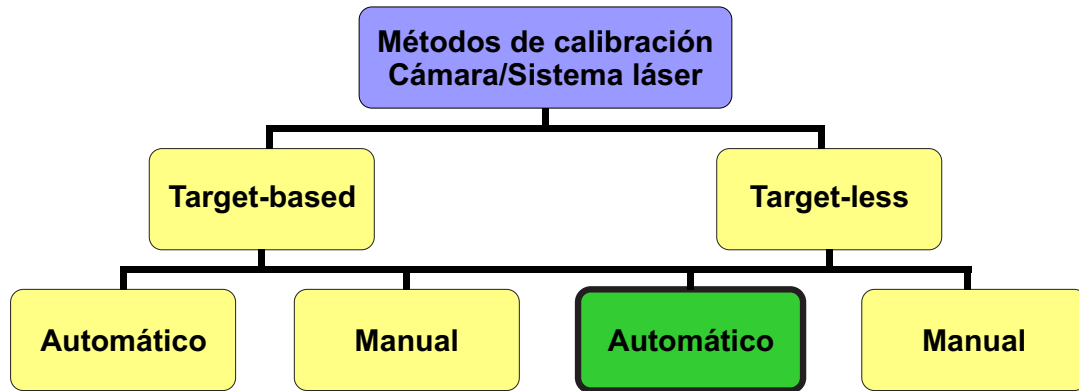


Figura 2.14: Clasificación de los métodos de calibración extrínseca entre sensores heterogéneos, adaptado de [21].

Capítulo 3: Metodología

3.1. Procedimiento de calibración en el sistema híbrido Estéreo–TVS

El presente capítulo expone de manera detallada la metodología desarrollada para la calibración de un sistema híbrido de percepción tridimensional, conformado por un Sistema de Visión Técnica (TVS) y un sistema de estereovisión. Dicha metodología integra procedimientos de calibración intrínseca y extrínseca, diseñados específicamente para minimizar la dependencia de patrones visuales estructurados, utilizando en su lugar superficies planas observadas simultáneamente por ambos subsistemas. El proceso se organiza en una serie de etapas secuenciales que incluyen: la caracterización angular de los motores del TVS, la evaluación de la inclinación y alineación de sus espejos, la definición de un método para la medición del *backlash* mecánico, y finalmente, la estimación de las transformaciones geométricas necesarias para alinear ambos marcos de referencia. Esta última se lleva a cabo mediante el ajuste de planos por mínimos cuadrados, empleando descomposición en valores singulares (SVD), lo que garantiza una integración precisa de las nubes de puntos generadas por cada subsistema.

El método de calibración propuesto plantea dos posibles configuraciones iniciales. La primera asume que las cámaras del sistema de visión estereo se calibran mediante un método convencional, como el planteado por Zhang [15], basado en la observación de un patrón planar —típicamente un tablero de ajedrez— desde múltiples vistas. Una vez estimados los parámetros intrínsecos y extrínsecos de las cámaras, el sistema estereo calibrado se emplea como instrumento auxiliar para la calibración intrínseca del Sistema de Visión Técnica (TVS).

La segunda configuración busca implementar un enfoque inverso: partiendo de un TVS previamente calibrado de forma manual, por ejemplo, utilizando objetos de geometría conocida dispuestos sobre una mesa óptica y entonces, se utiliza el TVS como elemento auxiliar para calibrar las cámaras. Este enfoque consiste en registrar, mediante el TVS, las coordenadas tridimensionales de una serie de

puntos sobre un objeto de referencia, mientras que simultáneamente se capturan imágenes de dichos puntos con las cámaras estéreo. Esta información permite establecer correspondencias 3D–2D entre los puntos medidos por el TVS y sus proyecciones en las imágenes, para así estimar los parámetros intrínsecos y extrínsecos de las cámaras y después realizar la calibración extrínseca del sistema estéreo. No obstante, en el presente trabajo se adopta exclusivamente el primer enfoque, en el que las cámaras se calibran con un método externo y se utiliza el sistema estéreo como instrumento auxiliar para la calibración del TVS. El segundo enfoque se plantea como una posibilidad complementaria para futuras investigaciones.

La Figura 3.1 presenta un esquema general del método propuesto, incluyendo ambos posibles inicios descritos. En las siguientes secciones se detallan las etapas específicas del método.

3.1.1. Parámetros intrínsecos del TVS

La estimación precisa de los parámetros intrínsecos del Sistema de Visión Técnica (TVS) resulta fundamental para asegurar su correcto funcionamiento y su alineación geométrica con el sistema estéreo. Estos parámetros tienen un impacto directo en la precisión de la reconstrucción 3D, así como en la correspondencia de puntos entre ambos sistemas. A continuación, se describe la importancia de los cuatro parámetros intrínsecos que fueron seleccionados en este estudio, considerando su gran influencia en el cálculo de coordenadas utilizando el método de triangulación dinámica:

- **Inclinación individual de los espejos en la apertura y el posicionador.** Este parámetro se refiere a la inclinación local de cada espejo respecto a su propio eje de rotación. Específicamente, se busca evitar errores de inclinación que provoquen que, al rotar, el haz láser describa una trayectoria no planar, elevándose o descendiendo indebidamente en lugar de mantener un desplazamiento estrictamente horizontal. Dichas inclinaciones inducen errores sistemáticos en la estimación de coordenadas, afectando la precisión de la triangulación dinámica y generando nubes de puntos distorsionadas.
- **Alineación entre los espejos del posicionador y la apertura** La alineación se refiere a que ambos espejos, considerados como un conjunto óptico, se encuentren correctamente

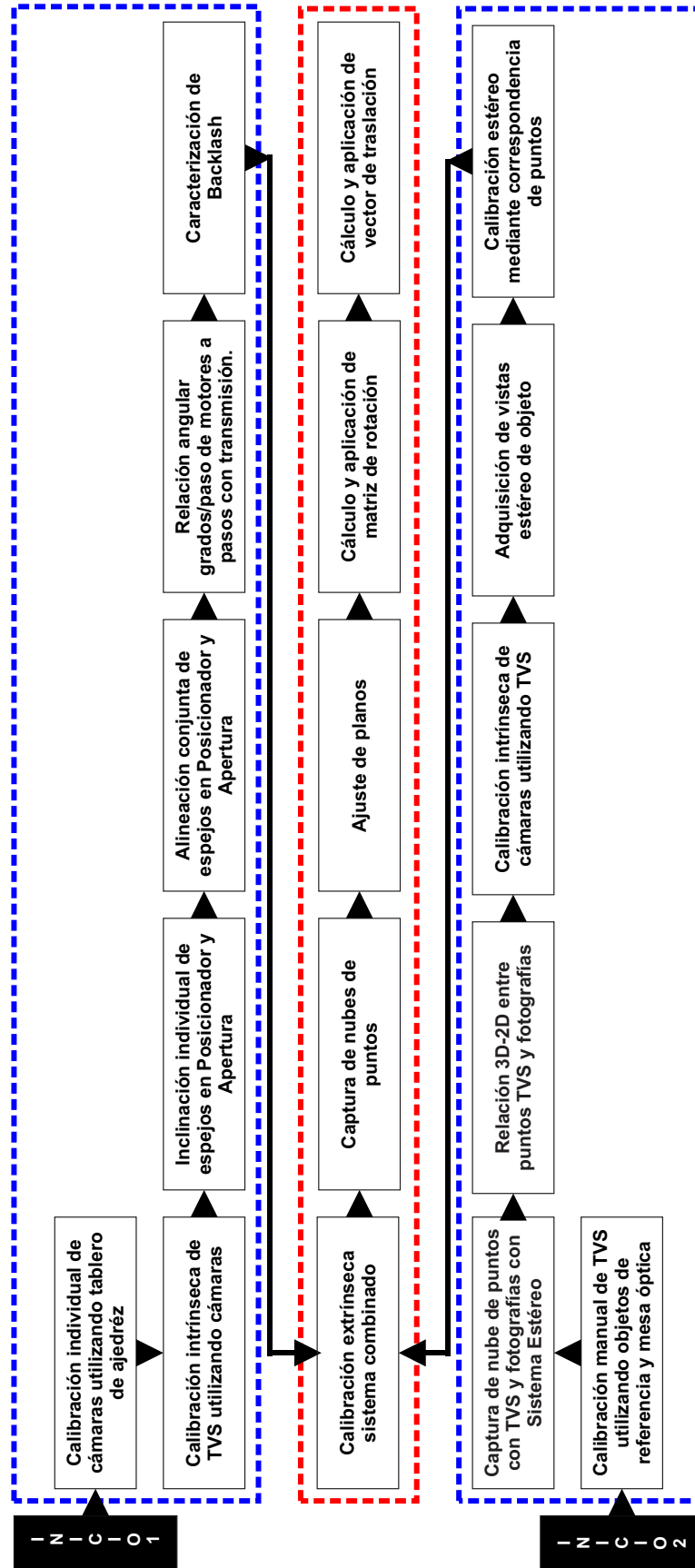


Figura 3.1: Diagrama general del método de calibración del sistema combinado Estéreo-TVS.

nivelados; es decir, que mantengan una altura relativa constante. Esta condición garantiza que la línea base del sistema esté bien definida, lo cual es esencial para la validez geométrica del modelo de triangulación dinámica. Una desalineación vertical entre los espejos puede dificultar la captura de la reflexión del láser a través de la apertura del TVS, introduciendo errores sistemáticos en la reconstrucción tridimensional.

- **Grados de avance por paso en el posicionador.** Este parámetro describe cuánto gira el espejo del elemento posicionador del sistema TVS por cada paso del motor, considerando también la transmisión mecánica. Es un valor clave para orientar con exactitud el haz láser en las direcciones horizontal y vertical. El ángulo que se genera en cada desplazamiento se usa directamente en los cálculos de coordenadas mediante triangulación dinámica, por lo que una estimación incorrecta puede afectar la precisión de la reconstrucción 3D, al alterar el valor angular con el que se proyecta el haz.
- **Backlash en el posicionador.** Debido a la presencia de transmisiones mecánicas, es necesario cuantificar el backlash, es decir, la holgura que aparece al invertir la dirección de movimiento durante un barrido. Si este error no se caracteriza ni compensa adecuadamente, puede acumularse a lo largo del recorrido, generando desviaciones en la ubicación de los puntos medidos y reduciendo la precisión de la reconstrucción tridimensional.

3.1.2. Calibración extrínseca del sistema Estéreo-TVS

Una vez definidos y calibrados los parámetros intrínsecos del TVS, se lleva a cabo la calibración extrínseca, cuyo objetivo es estimar la transformación rígida que alinea ambos subsistemas dentro de un marco de referencia común. El método propuesto en este trabajo consiste en adquirir simultáneamente nubes de puntos con cada sistema y ajustar un plano a cada una de ellas. A partir de dichos planos se construye un sistema de referencia local mediante vectores ortonormales —unitarios y mutuamente perpendiculares— definidos según la geometría del plano ajustado. Con estos marcos locales se calcula la matriz de rotación que alinea ambos sistemas, y posteriormente se determina la traslación a partir de la diferencia entre los centroides de las nubes de puntos.

3.2. Calibración intrínseca del TVS

La primera fase del procedimiento propuesto para la calibración del sistema híbrido Estéreo-TV_S requiere como condición previa la calibración intrínseca y extrínseca del sistema de estereovisión. Para ello, se adopta el método planteado por Zhang [15], el cual permite estimar los parámetros ópticos del modelo de cámara a partir de múltiples vistas de un patrón planar, típicamente un tablero de ajedrez, dispuesto en diferentes posiciones y orientaciones dentro del campo visual.

En la Figura 3.2 se muestra el patrón de ajedrez utilizado para esta etapa de calibración.

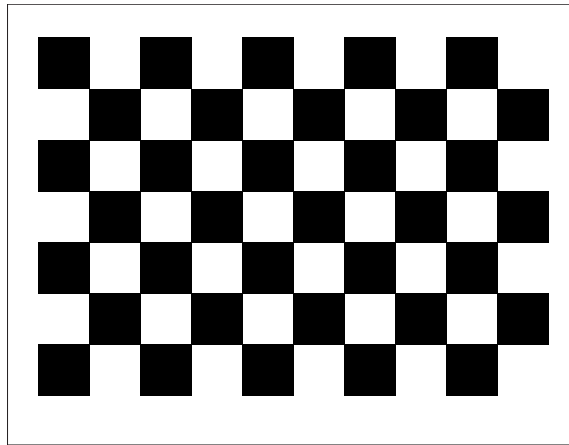


Figura 3.2: Patrón de ajedrez empleado en la calibración de cámaras.

La detección automática de esquinas del patrón se realiza mediante funciones de la biblioteca OpenCV en Python. Este procedimiento permite localizar con alta precisión los puntos de interés requeridos para la estimación de parámetros intrínsecos como la distancia focal, la posición del punto principal y los coeficientes de distorsión óptica. Posteriormente, se aplica una optimización global mediante *bundle adjustment* para refinar simultáneamente los parámetros intrínsecos de cada cámara y la geometría relativa entre ellas, obteniendo así la matriz de rotación $\mathbf{R}_{\text{stereo}}$ y el vector de traslación $\mathbf{t}_{\text{stereo}}$ que conforman la calibración extrínseca del sistema estéreo.

Una vez completada esta fase preliminar, se procede a la implementación del método desarrollado en este trabajo para la calibración del Sistema de Visión Técnica (TV_S). Este procedimiento se basa en la observación simultánea de una superficie plana visible por ambos sistemas, lo cual permite

establecer correspondencias geométricas sin necesidad de patrones estructurados.

La Figura 3.3 ilustra la disposición de dicha superficie, asumida como suficientemente plana, dentro del campo de visión común entre el sistema estéreo previamente calibrado y el TVS.



Figura 3.3: Ubicación de la superficie, que se asume como plana, dentro del campo de visión compartido por el sistema estéreo y el TVS.

La superficie empleada en el proceso no requiere poseer dimensiones conocidas ni presentar bordes visibles dentro del campo visual de los sensores. La única condición necesaria es que se trate de una superficie plana y que se encuentre simultáneamente dentro del campo de visión del sistema estéreo y del TVS. Asimismo, no es necesario conocer la distancia exacta a la que se encuentra respecto a ninguno de los sistemas, siempre que su observación conjunta sea posible. Para garantizar estabilidad y reproducibilidad en las mediciones, la superficie fue posicionada mediante soportes rígidos montados sobre una mesa óptica.

En las secciones siguientes se describe el procedimiento propuesto para la estimación de los parámetros intrínsecos del Sistema de Visión Técnica (TVS), seguido del proceso de calibración extrínseca que permite alinear espacialmente ambos subsistemas dentro del marco del sistema híbrido Estéreo-TV S.

3.2.1. Inclinación individual de los espejos del sistema TVS

El procedimiento propuesto permite estimar de forma simultánea la inclinación de los espejos montados en los dos mecanismos principales del TVS: el espejo del posicionador y el espejo ubicado en la apertura. Ambos elementos influyen en el cálculo de las coordenadas tridimensionales ya que corresponden a los ángulos de emisión y recepción, por lo que su inclinación relativa es crítica para la precisión del sistema.

La medición de la inclinación de los espejos se basa en la proyección del haz láser sobre una superficie plana y su observación mediante una o ambas cámaras del sistema estéreo. Aunque la triangulación tridimensional no es necesaria para este análisis, se puede realizar la captura simultánea con ambas cámaras como mecanismo de verificación y redundancia. El objetivo es estimar la inclinación de cada espejo a partir de la trayectoria proyectada bidimensionalmente del haz en el plano imagen.

El procedimiento seguido es el siguiente:

1. Se orienta el haz láser del TVS hacia una superficie plana colocada dentro del campo de visión del sistema estéreo.
2. Se captura una imagen rectificadas del haz utilizando una o ambas cámaras.
3. Se segmenta la región correspondiente a la mancha láser en el espacio HSV (de sus siglas en inglés: *Hue*, *Saturation*, *Value*) mediante umbrales que aíslan tonos correspondiente a la mancha del láser utilizado.
4. Se calcula el centroide del contorno más prominente utilizando los momentos espaciales de primer orden.
5. Se aplica una extensión vertical al contorno segmentado, basada en el análisis del canal de brillo (V), con el fin de mejorar la estimación del centroide en presencia de estructuras complejas o regiones superpuestas.
6. A partir del contorno extendido, se calcula un nuevo centroide refinado sobre la envolvente convexa de la región.

7. Se desplaza el haz una cantidad conocida de pasos mediante el motor del posicionador, y se repite el proceso para varias posiciones consecutivas.
8. Los centroides obtenidos se almacenan como un conjunto de puntos (x_i, y_i) en coordenadas de píxeles.
9. Se ajusta una recta a los puntos mediante regresión lineal utilizando la función `LinearRegression()` de `scikit-learn`.

El centroide de cada contorno segmentado se calcula a partir de los momentos espaciales M_{pq} , los cuales representan sumatorias ponderadas de las coordenadas x e y sobre los píxeles de la región binaria. En una imagen binaria $I(x, y)$, donde los píxeles pertenecientes al objeto tienen valor 1 y el fondo valor 0, el momento de orden (p, q) se define como:

$$M_{pq} = \sum_x \sum_y x^p y^q I(x, y) \quad (3.1)$$

En este contexto: - M_{00} representa el área del contorno (es decir, el número total de píxeles blancos),
 - M_{10} y M_{01} corresponden a los momentos de primer orden respecto a los ejes x e y , respectivamente.

El centroide (\bar{x}, \bar{y}) de la región se calcula como:

$$\bar{x} = \frac{M_{10}}{M_{00}}, \quad \bar{y} = \frac{M_{01}}{M_{00}} \quad (3.2)$$

Estos valores se obtienen directamente mediante la función `cv::moments()` de OpenCV, aplicada al contorno binarizado de la mancha láser.

Una vez determinados los centroides correspondientes a las distintas posiciones del haz, se ajusta una recta a este conjunto de puntos utilizando una regresión lineal, según la forma general mostrada en la Ecuación (3.3):

$$y = mx + b \quad (3.3)$$

A partir de la pendiente m , se calcula el ángulo de inclinación proyectada utilizando la Ecuación (3.4):

$$\theta = \arctan(m) \quad (3.4)$$

Evaluación de la inclinación del espejo en la apertura

El procedimiento para estimar la inclinación del espejo montado en la apertura es conceptualmente similar, con la diferencia de que este espejo está acoplado a un motor de corriente directa (C.D.) que gira de forma continua y no responde a comandos discretos por pasos.

En este caso, se proyecta un haz láser sobre el espejo de 45° en la apertura. Debido a la rotación continua del espejo, el haz describe una trayectoria rectilínea sobre la superficie plana. A partir de la imagen capturada, se detecta la línea formada por el barrido del haz y se ajusta una recta mediante el mismo procedimiento descrito previamente. La pendiente m de dicha recta se utiliza nuevamente para calcular la inclinación angular proyectada del espejo, mediante la Ecuación (3.4). En la Figura 3.4 se muestra una comparación visual entre espejos sin inclinación y aquellos que presentan un ángulo de inclinación no deseado.

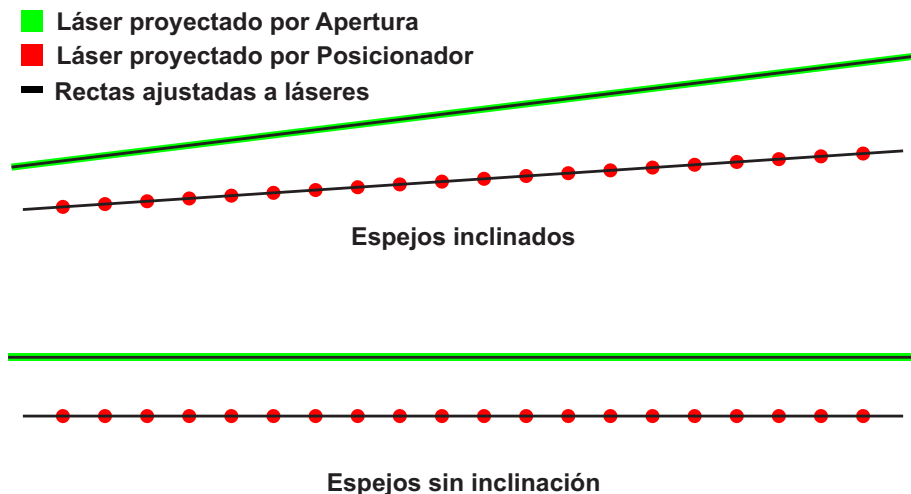


Figura 3.4: Comparación entre espejos inclinados y no inclinados.

En caso de detectarse una inclinación no deseada, los espejos del sistema requieren ajustes mecánicos

para restablecer su correcta orientación. Sin embargo, como línea de investigación futura se propone implementar una compensación vía software para el espejo del posicionador, aprovechando que su ángulo de rotación es conocido en todo momento. Esto no aplica en el caso del espejo ubicado en la apertura debido a que su posición angular no puede determinarse con precisión durante la rotación continua. En consecuencia, cualquier corrección debe realizarse exclusivamente mediante ajustes mecánicos en el montaje.

3.2.2. Alineación entre los espejos del posicionador y la apertura

Una vez verificada la inclinación individual de los espejos, se procede a evaluar su alineación conjunta, entendida como la consistencia en altura entre ambos elementos ópticos. Esta evaluación es crucial, ya que los dos espejos definen conjuntamente la base del triángulo geométrico empleado en el cálculo de coordenadas tridimensionales. Cualquier desajuste relativo puede introducir errores sistemáticos en la estimación de profundidad.

Para esta evaluación se utilizan las mismas imágenes obtenidas durante el procedimiento de inclinación individual. Si los espejos se encuentran correctamente coalineados, las trayectorias proyectadas del haz —una generada por el espejo del posicionador y otra por el espejo de la apertura— deben ser coincidentes. En particular, las rectas ajustadas a partir de los centroides obtenidos deben presentar valores similares tanto de pendiente m como de ordenada al origen b , de acuerdo con la Ecuación (3.3).

Una diferencia significativa entre los parámetros m o b sugiere una desalineación relativa, que puede deberse, por ejemplo, a una discrepancia en la altura de montaje de los elementos ópticos. En tales casos, se requiere un ajuste mecánico en la posición o inclinación de alguno de los espejos hasta lograr la coincidencia entre ambas trayectorias.

En la Figura 3.5 se muestra un caso ideal en el que ambos espejos se encuentran correctamente coalineados. En este escenario, las trayectorias proyectadas del haz láser generadas por el posicionador (en rojo) y por la apertura (en verde) coinciden a lo largo del eje horizontal. Las rectas ajustadas a cada conjunto de puntos (líneas negras) también se superponen, evidenciando la consistencia

angular y posicional entre ambos elementos ópticos.

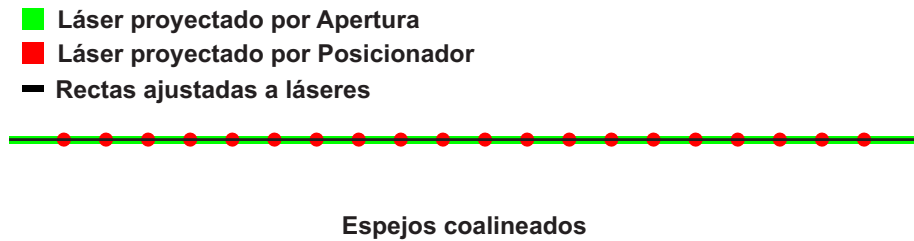


Figura 3.5: Comparación entre espejos de apertura y posicionador coalineados.

3.2.3. Caracterización angular de los motores: relación grados por paso

Una vez verificada la coalineación de los espejos del TVS, se procede a caracterizar el comportamiento angular de los motores a pasos que controlan el posicionamiento del haz láser en los ejes horizontal y vertical del sistema. Esta caracterización es necesaria para establecer la relación entre el número de pasos comandados al motor y el desplazamiento angular efectivo del haz, así como para detectar la presencia de efectos como el backlash, inducido por las transmisiones mecánicas acopladas al actuador.

Procedimiento para estimar la relación grados por paso

El presente procedimiento permite estimar la relación entre el número de pasos comandados al motor del posicionador y el desplazamiento angular efectivo del haz láser. Para ello, se registran las coordenadas tridimensionales del punto de impacto del haz sobre una superficie plana antes y después de aplicar un desplazamiento controlado. A partir de estos puntos, se calcula la proyección angular del haz en cada caso, y se determina su inclinación inicial y final. La diferencia entre ambos ángulos, dividida por el número de pasos aplicados, permite obtener la relación buscada en unidades de *grados/paso*. A continuación se detalla el procedimiento paso a paso.

1. Se posiciona el haz láser del TVS sobre una superficie plana ubicada a una distancia arbitraria d_1 , como se mostró en la Figura 3.3.
2. Se adquieren imágenes del haz utilizando ambas cámaras del sistema estéreo, se rectifican y

se procesa la región de interés en el espacio HSV para detectar el contorno de la mancha.

3. A partir del contorno se calcula el centroide de la mancha del láser $C_l = (x_c, y_c)$, y se triangulan las coordenadas tridimensionales del punto de impacto: $P_1 = (X_1, Y_1, Z_1)$.
4. Se traslada la superficie plana a una nueva posición d_2 , sin necesidad de conocer esta distancia con precisión, y se repite el proceso para obtener un segundo punto: $P_2 = (X_2, Y_2, Z_2)$.
5. A continuación, se ordena un desplazamiento horizontal del haz hacia la izquierda mediante un número conocido de pasos N_{pasos} , y se registra un nuevo punto P_3 en la posición d_2 , y luego otro punto P_4 al retornar la superficie a la posición d_1 .

De este modo, se obtienen dos pares de puntos: P_1, P_2 antes del movimiento del motor, y P_3, P_4 después. Cada par define una dirección del haz correspondiente a una orientación angular distinta del espejo del posicionador.

Cálculo del ángulo de emisión y estimación de la relación

Cada par de puntos triangulados define una trayectoria proyectada del haz láser en el plano XY , donde el eje X representa la profundidad (distancia desde el sistema de cámaras) y el eje Y corresponde al desplazamiento horizontal. De este modo, la dirección angular proyectada del haz puede calcularse a partir de la razón entre el desplazamiento horizontal y la profundidad, según se muestra en la Ecuación (3.5):

$$\theta_i = \tan^{-1} \left(\frac{Y_2 - Y_1}{X_2 - X_1} \right), \quad \theta_f = \tan^{-1} \left(\frac{Y_3 - Y_4}{X_3 - X_4} \right) \quad (3.5)$$

La Figura 3.6 ilustra los triángulos rectángulos formados por los pares de puntos antes y después del desplazamiento angular. Los catetos ΔY y ΔX corresponden al desplazamiento horizontal del haz y a la variación de profundidad entre los puntos observados, respectivamente.

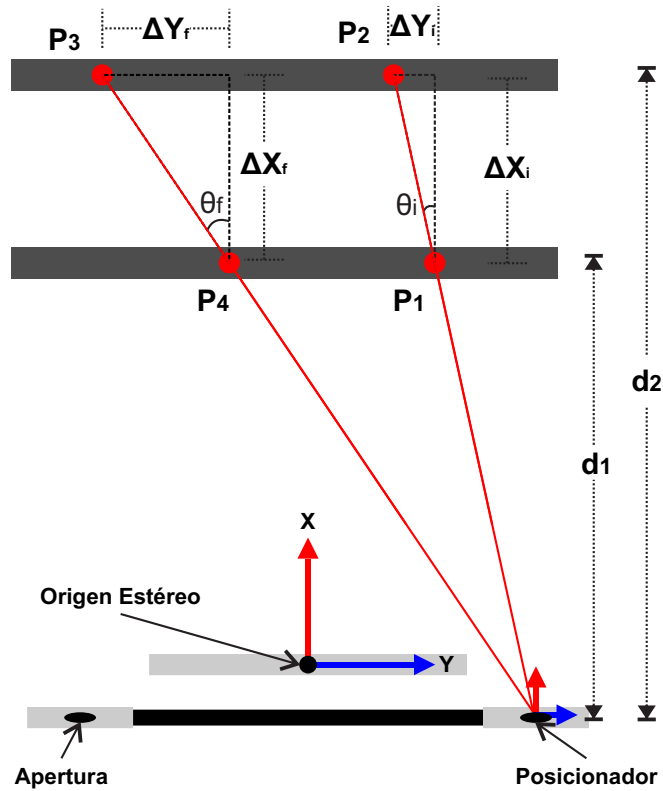


Figura 3.6: Triángulos proyectados en el plano XY a partir de los puntos triangulados antes y después del desplazamiento del motor.

Conociendo la diferencia angular inducida por el movimiento del motor y el número exacto de pasos aplicados se calcula la relación angular buscada en unidades de grados/paso, como se indica en la Ecuación (3.6):

$$\frac{\text{grados}}{\text{paso}} = \frac{\theta_f - \theta_i}{N_{\text{pasos}}} \quad (3.6)$$

Este procedimiento se repite para los cuatro sentidos de movimiento del posicionador (izquierda, derecha, arriba y abajo), ya que el sistema de transmisión puede presentar comportamientos distintos en cada dirección debido a tolerancias mecánicas.

3.2.4. Método propuesto para medición del backlash

El *backlash* o juego mecánico constituye una forma de histéresis angular que se manifiesta en sistemas con transmisión indirecta, como motores paso a paso acoplados mediante engranajes o

mecanismos flexibles. Este fenómeno se presenta al invertir el sentido de giro del actuador, generando una zona muerta angular en la cual los primeros pasos comandados no producen desplazamiento efectivo. En el contexto del direccionamiento del haz láser en el Sistema de Visión Técnica (TVS), este efecto puede inducir errores sistemáticos en la estimación de coordenadas tridimensionales, por lo que resulta necesario cuantificarlo y compensarlo durante el escaneo.

Procedimiento experimental básico

Para medir el backlash se emplea una superficie plana fija, sobre la cual se proyecta el haz láser. La posición del punto de impacto se determina mediante triangulación estéreo. El procedimiento mínimo requiere tres capturas consecutivas del punto láser, obtenidas al realizar un movimiento de ida y retorno controlado sobre un mismo eje. La secuencia experimental propuesta se describe a continuación:

1. Posicionar el haz láser sobre la superficie y registrar el punto inicial $P_1 = (X_1, Y_1, Z_1)$ mediante triangulación estéreo.
2. Aplicar un desplazamiento angular hacia un sentido (por ejemplo, izquierda) mediante N_{pasos} del motor, y registrar el nuevo punto P_2 .
3. Invertir el sentido del motor y aplicar nuevamente N_{pasos} , obteniendo un tercer punto P_3 .

En condiciones ideales (sin backlash), el punto P_3 coincidiría con P_1 . Sin embargo, en presencia de juego mecánico, el espejo del posicionador no recupera su orientación original, y se produce una discrepancia horizontal proyectada entre ambas posiciones.

Estimación del ángulo de backlash

La magnitud del backlash se estima a partir de la diferencia en coordenadas Y entre el punto inicial y el de retorno, según se muestra en la Ecuación (3.7):

$$d_b = Y_3 - Y_1 \tag{3.7}$$

Esta distancia d_b representa el cateto opuesto de un triángulo rectángulo, cuyo cateto adyacente es la distancia entre el posicionador y la superficie de proyección, denotada como D_{TVS} , que puede conocerse a partir del experimento anteriormente descrito para la caracterización de grados por paso, siendo aquella distancia a la cual convergen las dos rectas trazadas sobre el plano XY o XZ según corresponda al backlash horizontal o vertical. A partir de esta geometría, se calcula el ángulo de backlash asociado al sentido de retorno utilizando la función tangente inversa, como se indica en la Ecuación (3.8):

$$\theta_b = \tan^{-1} \left(\frac{d_b}{D_{TVS}} \right) \quad (3.8)$$

Este ángulo cuantifica el desplazamiento angular que no fue efectivamente ejecutado durante el movimiento de retorno, debido a la histéresis de la transmisión mecánica.

Estrategia de medición bidireccional

Para mejorar la eficiencia del procedimiento y obtener información en ambos sentidos de desplazamiento, se propone un esquema iterativo que alterna movimientos de ida y vuelta. En cada iteración se registra el punto inicial, se aplica un desplazamiento en una dirección y luego el mismo número de pasos en dirección opuesta. La discrepancia entre la primera y la tercera posición corresponde al backlash del segundo movimiento, es decir, del sentido de retorno.

Interpretación direccional:

- Si el ciclo comienza hacia la izquierda y luego retorna a la derecha, la discrepancia observada en P_3 respecto a P_1 refleja el backlash asociado al movimiento hacia la derecha.
- Si se invierte el orden (primero derecha, luego izquierda), se obtiene el backlash correspondiente al movimiento hacia la izquierda.

Este procedimiento también se replica para el eje vertical, permitiendo estimar el backlash superior e inferior mediante secuencias análogas de movimientos arriba-abajo.

3.3. Calibración extrínseca en mesa óptica

Después de haber ajustado los parámetros intrínsecos del TVS. La segunda etapa del método desarrollado en este estudio consiste en la calibración extrínseca del sistema híbrido Estéreo-TVS. Dado que cada subsistema opera en su propio marco de referencia, es necesario determinar una transformación geométrica que incluye traslación y rotación espacial para alinear ambos sistemas de coordenadas. Al realizar este proceso se logra integrar los datos adquiridos de manera independiente en una única nube de puntos tridimensional precisa y coherente.

Las diferentes etapas del procedimiento de calibración extrínseca se ilustran En la Figura 3.7. En este diagrama, las actividades relacionadas con la adquisición de datos se identifican mediante una línea de contorno color café, mientras que los pasos vinculados con la traslación de coordenadas están delimitados por una línea de contorno azul. Por último, las operaciones relacionadas con la rotación espacial se distinguen por un marco de color anaranjado.

En las subsecciones siguientes, se presentan de manera detallada cada una de estas tres fases del procedimiento.

3.3.1. Adquisición de datos y generación de nubes de puntos

La adquisición de datos constituye la etapa inicial del procedimiento de calibración extrínseca, en esta etapa se capturan y registran simultáneamente las coordenadas tridimensionales de un conjunto de puntos ubicados sobre la superficie de interés, la cual debe ser plana. Para ello, se emplean dos sistemas de visión independientes: el sistema de estereovisión y el sistema TVS.

Cada nube de puntos obtenida se expresa en el marco de referencia del sistema que la capturó. Cuando el campo de visión (FoV) y el rango de medición de ambos sensores se superponen, es posible identificar puntos homólogos observados desde distintas perspectivas.

El proceso inicia con la definición del número total de puntos a medir, representado por el parámetro N . Una vez establecido este valor, el sistema TVS realiza un barrido sistemático sobre la superficie, y obtiene las coordenadas tridimensionales mediante su método de triangulación dinámica, definido

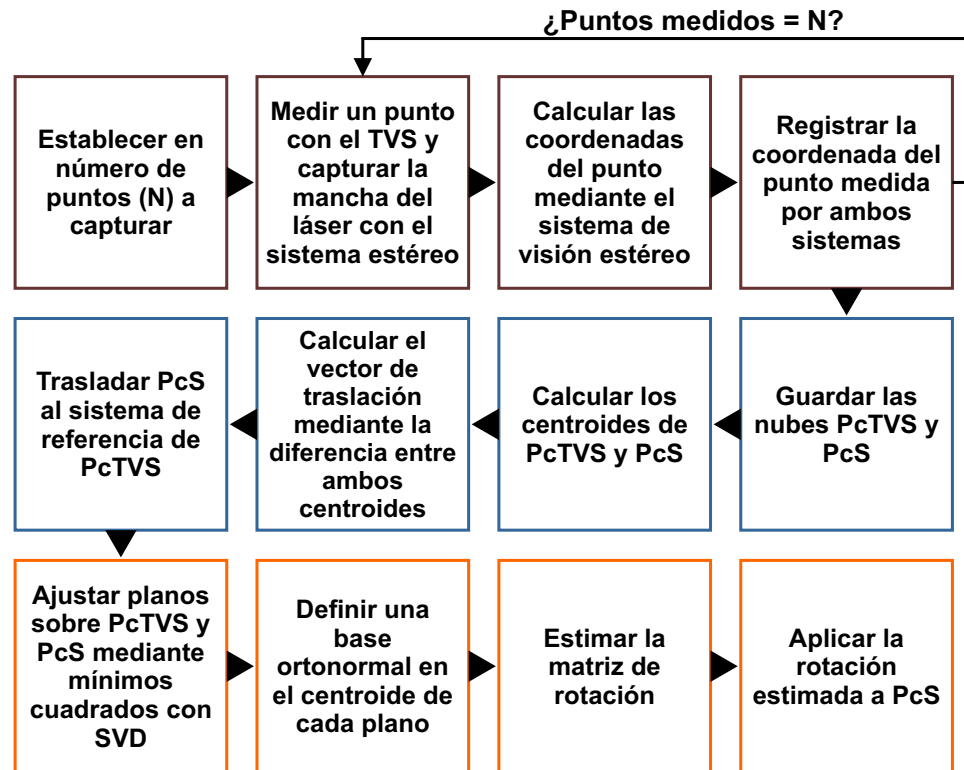


Figura 3.7: Procedimiento para la adquisición de datos y determinación de parámetros extrínsecos.

por la distancia entre dos espejos y los ángulos de posicionamiento y apertura de estos. Durante el escaneo, el TVS registra en cada posición i , con $i = 1, 2, \dots, N$, la coordenada tridimensional del punto correspondiente, denotada como P_{cTVS} , expresada en su propio marco de referencia.

De forma paralela, el sistema de estereovisión captura imágenes del punto láser utilizando sus dos cámaras en cada una de las posiciones donde el TVS realizó una medición. Posteriormente, se aplican técnicas de procesamiento digital a cada par de imágenes adquiridas, incluyendo nuevamente segmentación mediante umbralización y detección de contornos.

La umbralización permite separar los píxeles, correspondientes al láser, de aquellos asociados al fondo. Posteriormente, se localiza el centroide del punto láser en ambas imágenes, definido como (c_x^L, c_y^L) en la imagen izquierda y (c_x^R, c_y^R) en la imagen derecha. A partir de estos centroides, se emplea triangulación estéreo para estimar las coordenadas tridimensionales del punto observado, denotado como $P_{cS,i}$, el cual se encuentra en el marco de referencia del sistema de estereovisión.

Este procedimiento se repite iterativamente hasta completar la adquisición de N puntos, generando

dos nubes tridimensionales independientes, como se describe en las Ecuaciones (3.9) y (3.10):

$$P_{cTVS} = \{P_{TVS,1}, P_{TVS,2}, \dots, P_{TVS,N}\} \quad (3.9)$$

$$P_{cS} = \{P_{S,1}, P_{S,2}, \dots, P_{S,N}\} \quad (3.10)$$

En la Figura 3.8 se muestra una representación de las nubes de puntos adquiridas por ambos sistemas.

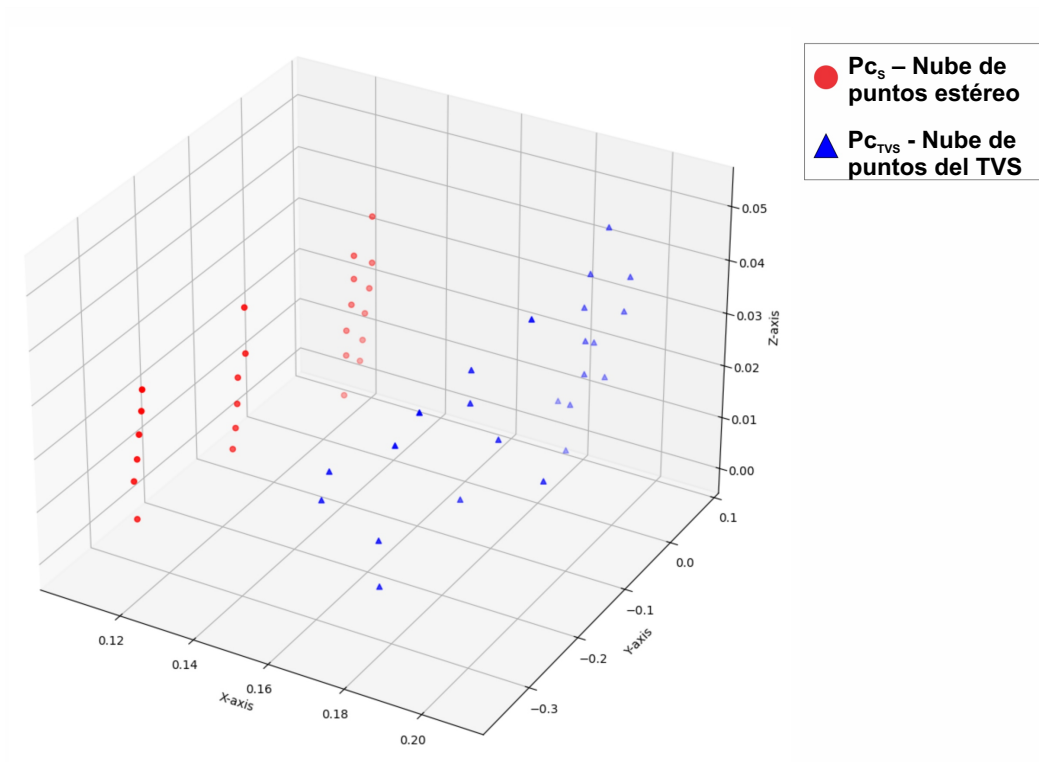


Figura 3.8: Comparativa entre las nubes de puntos generadas por los sistemas TVS y estéreo.

3.3.2. Traslación de nubes de puntos

Como sabemos, dado que cada sistema opera en un marco de referencia distinto, es necesario aplicar una transformación geométrica para alinear ambas nubes de puntos en un sistema de referencia común. En esta sección abordaremos el proceso para la estimación de la traslación necesaria que

permite ubicar las nubes dentro de un único marco de referencia común.

De manera general, una nube de puntos tridimensional P puede definirse formalmente mediante el conjunto que se muestra en la Ecuación (3.11):

$$P = \{(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)\} \quad (3.11)$$

Podemos determinar el centroide \mathbf{C} de dicha nube, el cual se calcula como el promedio aritmético de las coordenadas espaciales de todos los puntos que la conforman, como se indica en la Ecuación (3.12):

$$\mathbf{C} = \left(\frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n y_i, \frac{1}{n} \sum_{i=1}^n z_i \right) \quad (3.12)$$

Para trasladar la nube P_{cS} de tal forma que su centroide coincida con el de la nube P_{cTVS} , se determina el vector de traslación como se muestra en la Ecuación (3.13):

$$V_t = C_{TVS} - C_S \quad (3.13)$$

Este vector se aplica sobre cada punto perteneciente a la nube P_{cS} , generando así una nueva nube trasladada P'_{cS} , la cual queda definida por la Ecuación (3.14):

$$P'_{cS} = \{p_i + V_t \mid p_i \in P_{cS}\} \quad (3.14)$$

La Figura 3.9 muestra de manera gráfica este procedimiento, ilustrando cómo la diferencia entre los centroides de la nube de referencia y la nube objetivo define el vector de traslación \mathbf{V}_t , necesario para integrar ambas estructuras en un mismo sistema de coordenadas.

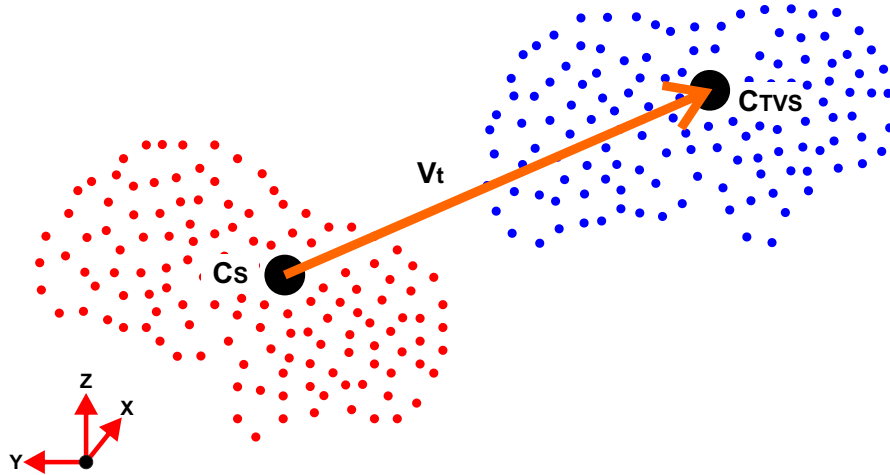


Figura 3.9: Representación esquemática del vector de traslación \mathbf{V}_t , definido como la diferencia entre los centroides de las nubes de puntos capturadas por el sistema estéreo y el sistema TVS.

3.3.3. Rotación mediante ajuste de planos

Para completar la alineación de las nubes de puntos tridimensionales, además de la traslación, es necesario aplicar una rotación. Para ello, en este método se propone el ajuste de un plano utilizando el método de mínimos cuadrados (LSQ) por descomposición en valores singulares (SVD) a ambas nubes de puntos.

Ajuste de planos

El ajuste de un plano a una nube de puntos tridimensional implica asumir que dicha nube $P = \{p_i\}$ se aproxima a una superficie plana. El objetivo es determinar tanto un vector normal que defina la orientación del plano, como un punto de referencia que minimice la distancia entre los puntos y dicho plano. Para ello, se emplea el centroide como referencia, dado que representa el punto que minimiza la suma de las distancias cuadráticas respecto al conjunto de puntos, sirviendo como un punto central equilibrado dentro de la nube.

Inicialmente, se determina el centroide de la nube mediante el cálculo del promedio de las coordenadas tridimensionales de cada punto, como se mostró previamente en la Ecuación (3.12).

Posteriormente, se trasladan los puntos de la nube a un sistema de referencia centrado en dicho

centroide, como se indica en la Ecuación (3.15):

$$p'_i = p_i - c \quad (3.15)$$

Una vez centrada la nube, se construye la matriz A que contiene los vectores p'_i como filas, y se aplica sobre ella la descomposición en valores singulares (SVD), según la Ecuación (3.16):

$$A = U\Sigma V^T \quad (3.16)$$

El último vector columna de V corresponde al vector normal n que define el plano ajustado, al minimizar la distancia cuadrática entre los puntos y dicho plano.

La ecuación general del plano estimado a partir de este vector normal se expresa en la Ecuación (3.17):

$$ax + by + cz + d = 0 \quad (3.17)$$

En esta expresión, a, b y c son las componentes del vector normal, mientras que el término d se calcula a partir del producto escalar entre el vector normal y el centroide.

Este procedimiento se aplica por separado a las dos nubes generadas por los sistemas TVS y estéreo, obteniendo dos planos ajustados que denominamos π_{TVS} y π_S , respectivamente. La Figura 3.10 muestra estos planos ajustados sobre las nubes de puntos correspondientes.

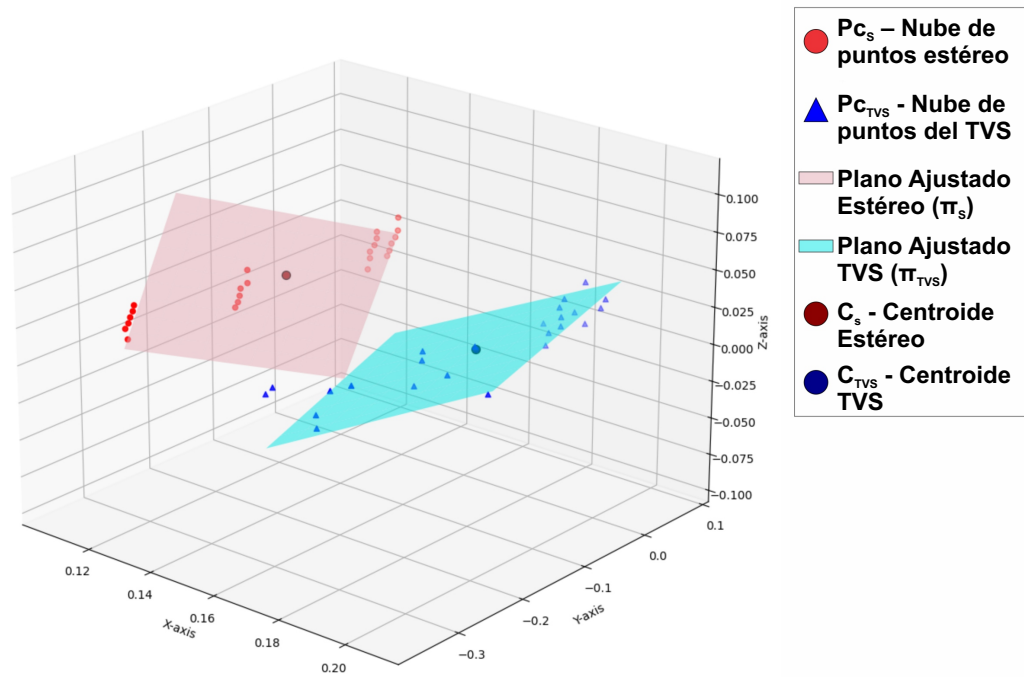


Figura 3.10: Planos ajustados a partir de las nubes de puntos obtenidas por los sistemas TVS y estéreo.

Definición de un sistema ortonormal asociado al plano

A partir de la ecuación general del plano, se define un sistema de referencia ortonormal compuesto por tres vectores: el vector normal \mathbf{z} , y dos vectores ortogonales contenidos en el plano, \mathbf{x} y \mathbf{y} .

El vector normal \mathbf{z} se obtiene normalizando el vector de coeficientes (a, b, c) , como se indica en la Ecuación (3.18):

$$\mathbf{z} = \frac{1}{\sqrt{a^2 + b^2 + c^2}} \begin{bmatrix} a & b & c \end{bmatrix} \quad (3.18)$$

A continuación, se define un vector \mathbf{x} contenido en el plano, mediante la ortogonalización de un vector inicial $\mathbf{v}_{inicial}$, como se muestra en la Ecuación (3.19):

$$\mathbf{x} = \mathbf{v}_{inicial} - (\mathbf{v}_{inicial} \cdot \mathbf{z})\mathbf{z} \quad (3.19)$$

El tercer vector, \mathbf{y} , se obtiene como el producto vectorial entre \mathbf{z} y \mathbf{x} , garantizando ortonormalidad.

Con estos vectores se definen los marcos locales de cada plano ajustado, según la Ecuación (3.20):

$$F_{TVS} = [\mathbf{x}_{TVS}, \mathbf{y}_{TVS}, \mathbf{z}_{TVS}], \quad F_S = [\mathbf{x}_S, \mathbf{y}_S, \mathbf{z}_S] \quad (3.20)$$

La Figura 3.11 muestra la visualización conjunta de estos marcos aplicados a sus respectivas nubes.

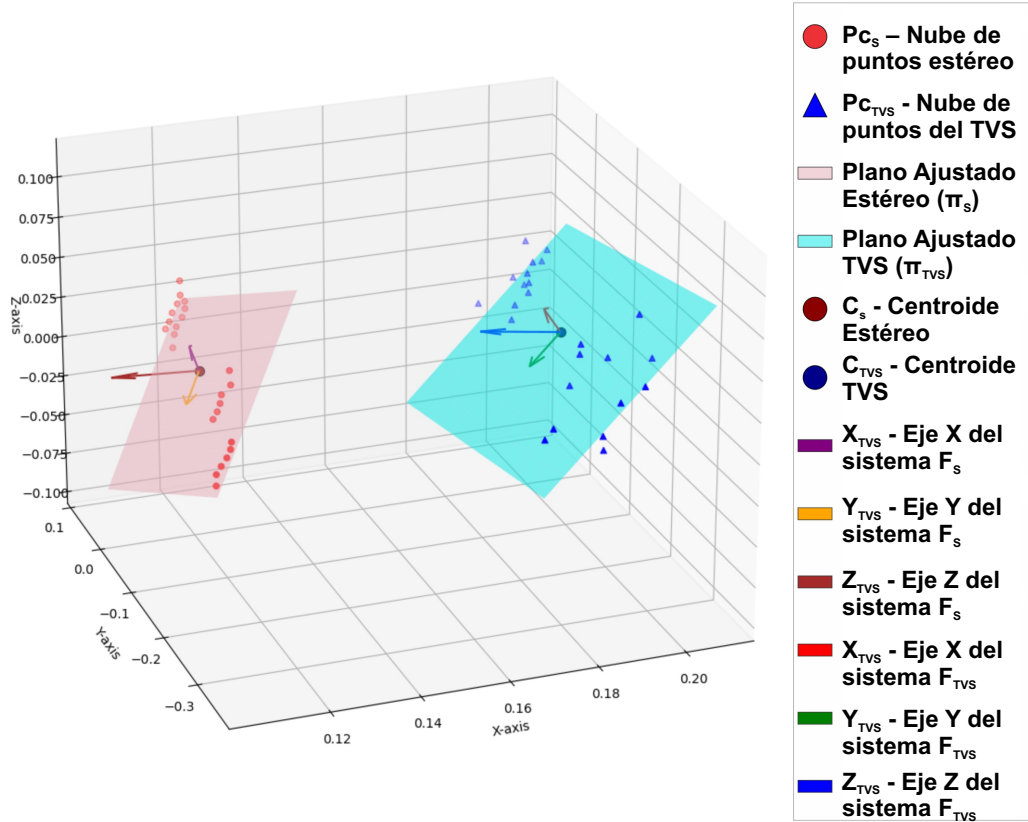


Figura 3.11: Comparación entre las nubes de puntos del sistema estéreo y TVS, incluyendo planos ajustados, centroides y vectores ortonormales definidos.

Estimación de la matriz de rotación

Las orientaciones de ambos marcos locales pueden representarse matricialmente como se muestra en las ecuaciones (3.21) y (3.22) :

$$F_{TVS} = \begin{bmatrix} \mathbf{x}_{TVS} & \mathbf{y}_{TVS} & \mathbf{z}_{TVS} \end{bmatrix} \quad (3.21)$$

$$F_S = \begin{bmatrix} \mathbf{x}_S & \mathbf{y}_S & \mathbf{z}_S \end{bmatrix} \quad (3.22)$$

A partir de estas matrices, la rotación requerida para alinear ambos marcos se calcula como:

$$R = F_{TVS} F_S^T \quad (3.23)$$

Rotación de la nube de puntos

Dado que la rotación se realiza respecto al origen, primero es necesario centrar la nube P_{cS} trasladando su centroide al origen, como se muestra en la Ecuación (3.24):

$$P_{cS}^{\text{centrado}} = P_{cS} - c_S \quad (3.24)$$

Luego, se aplica la rotación calculada sobre la nube centrada, como se indica en la Ecuación (3.25):

$$P_{cS}^{\text{rotado}} = R \cdot P_{cS}^{\text{centrado}} \quad (3.25)$$

Del mismo modo, se aplica la rotación a los vectores ortonormales del marco F_S , como se muestra en las Ecuaciones (3.26) y (3.27):

$$F_S^{\text{rotado}} = R \cdot F_S = \begin{bmatrix} \mathbf{x}_S^{\text{rotado}} & \mathbf{y}_S^{\text{rotado}} & \mathbf{z}_S^{\text{rotado}} \end{bmatrix} \quad (3.26)$$

$$\begin{bmatrix} \mathbf{x}_S^{\text{rotado}} \\ \mathbf{y}_S^{\text{rotado}} \\ \mathbf{z}_S^{\text{rotado}} \end{bmatrix} = R \cdot \begin{bmatrix} \mathbf{x}_S \\ \mathbf{y}_S \\ \mathbf{z}_S \end{bmatrix} \quad (3.27)$$

Traslación final para coincidencia de centroides

Para finalizar el alineamiento, se aplica una traslación adicional que posiciona el centroide de la nube rotada c_S^{rotado} sobre el centroide de referencia c_{TVS} , como se indica en la Ecuación (3.28):

$$t = c_{TVS} - c_S^{\text{rotado}} \quad (3.28)$$

La nube transformada final se obtiene al aplicar dicha traslación:

$$P_{cS}^{\text{transformado}} = P_{cS}^{\text{rotado}} + t \quad (3.29)$$

La Figura 3.12 muestra el resultado final del proceso de alineamiento geométrico entre las nubes de puntos.

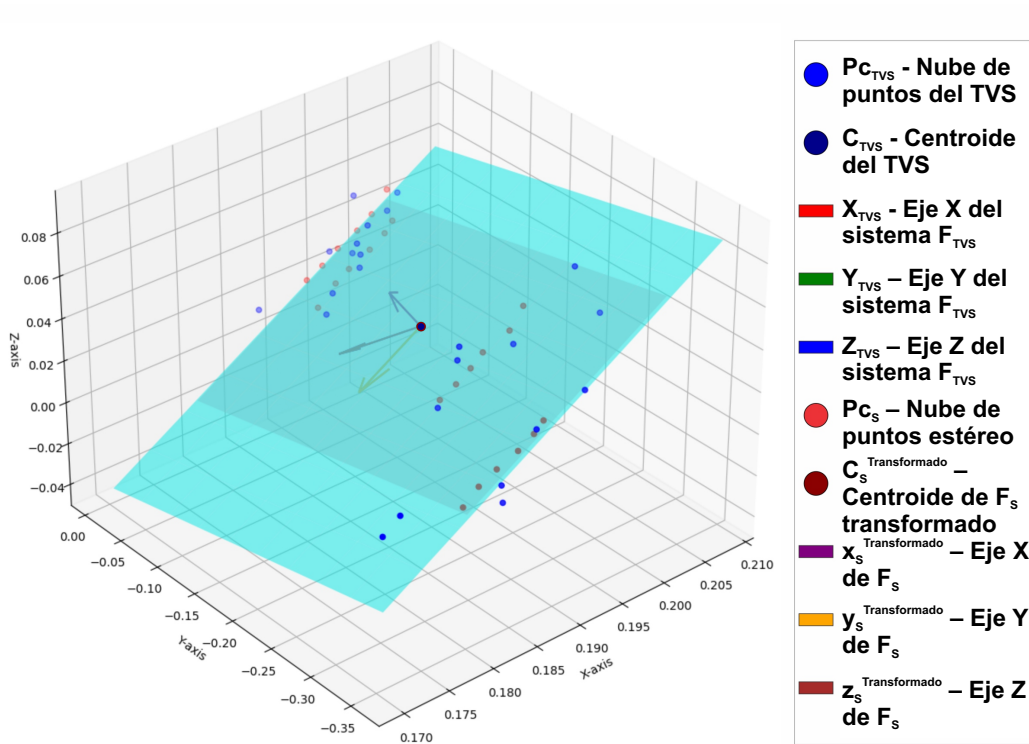


Figura 3.12: Resultado final del alineamiento entre las nubes de puntos TVS y estéreo tras la aplicación de transformaciones geométricas.

3.4. Ejemplo de aplicación: calibración extrínseca montando sistemas en robot con movimiento intermitente

En la sección previa, la calibración extrínseca del sistema combinado Estéreo-TVS se realizó empleando una mesa óptica. En cambio, en esta sección se explora cómo, con ciertas modificaciones, dicha calibración extrínseca puede efectuarse directamente en aplicaciones donde se realizan escaneos de superficies planas, por ejemplo en conductos de ventilación.

Aunque habitualmente la alineación entre los sistemas se efectúa previamente al escaneo, la metodología propuesta aquí permite estimar la posición relativa entre los sistemas utilizando como referencia las paredes planas del propio conducto. Esta capacidad es especialmente útil cuando el robot móvil se encuentra en lugares inaccesibles, en condiciones de completa oscuridad o cuando movimientos repentinos alteran la alineación original entre los sistemas de visión artificial.

A continuación, se describen las diferencias necesarias en método propuesto, con respecto al procedimiento presentado anteriormente, realizado sobre una mesa óptica. En la Figura 3.13 se muestra el esquema completo del método, abarcando desde la adquisición inicial de datos y su procesamiento, hasta la estimación y posterior aplicación de transformaciones geométricas mediante traslación y rotación.

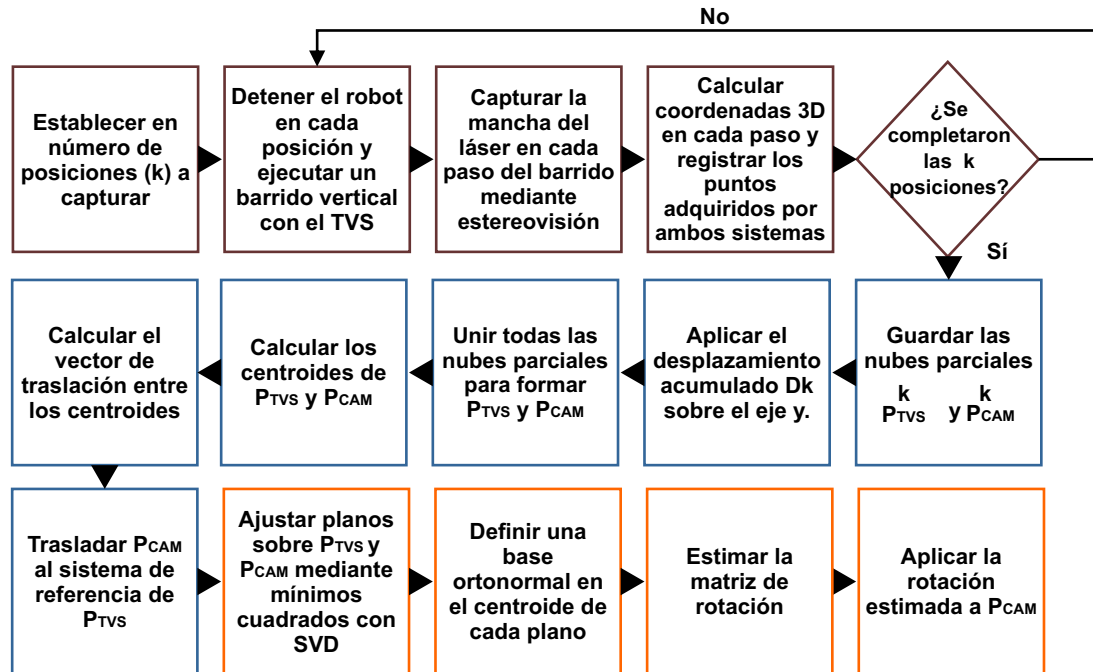


Figura 3.13: Esquema metodológico para la estimación de parámetros extrínsecos en un sistema Estéreo-TVIS montado sobre un robot móvil.

Se puede observar que gran parte del método permanece inalterada, con las principales diferencias concentradas en la etapa de adquisición de datos y la generación de nubes de puntos. En este enfoque, que emplea un robot móvil, la captura de datos consiste en registrar simultáneamente las mediciones del TVS y las imágenes de la proyección láser mediante el sistema de estereovisión, mientras el dispositivo avanza progresivamente a lo largo del eje y del marco de referencia del robot. En cada posición específica k , el robot se detiene momentáneamente para registrar la información, lo que permite la adquisición de conjuntos de puntos más extensos a medida que aumenta el número de posiciones consideradas. Antes de iniciar el experimento, se define previamente el número total de posiciones a capturar.

Debemos considerar que cada conjunto de datos capturado debe ajustarse teniendo en cuenta el desplazamiento progresivo del robot móvil. El desplazamiento acumulado en la posición k , denotado como D_k , se calcula como la suma de los desplazamientos individuales registrados hasta esa posición, tal como se expresa en la Ecuación (3.30).

$$D_k = \sum_{j=1}^k d_j \quad (3.30)$$

Finalmente, las nubes de puntos resultantes de ambos sistemas (TVS y sistema estéreo) se generan uniendo secuencialmente los puntos adquiridos en cada posición k . La construcción completa de cada nube tridimensional se expresa formalmente en las Ecuaciones (3.31) y (3.32):

$$P_{cTVS} = \bigcup_{k=1}^m \{(x_i^k, y_i^k, z_i^k)\}_{i=1}^N \quad (3.31)$$

$$P_{cS} = \bigcup_{k=1}^m \{(x_i^k, \tilde{y}_i^k, z_i^k)\}_{i=1}^N, \quad \text{con} \quad \tilde{y}_i^k = y_i^k + D_k \quad (3.32)$$

donde $i = 1, 2, \dots, N$ corresponde al índice de cada punto en una captura específica, N representa la cantidad total de puntos adquiridos por posición, y $\tilde{y}_i^k = y_i^k + D_k$ indica el ajuste por desplazamiento acumulado mientras el robot avanza.

Después de la etapa de adquisición de datos, el resto del método permanece inalterado. La nube de puntos obtenida mediante el sistema estéreo se transforma al marco de referencia del TVS utilizando un vector de traslación, calculado a partir de la diferencia entre los centroides de ambas nubes de puntos. Posteriormente, se ajusta un plano a cada nube de puntos, lo que permite definir un sistema de referencia asociado a cada plano. A partir de estos sistemas, se determina la matriz de rotación que alinea ambos planos, la cual, al aplicarse a los datos de la nube de puntos del sistema de estereovisión, permite su correcta transformación.

Capítulo 4: Experimentación y resultados

Este capítulo describe la experimentación realizada para validar el procedimiento de calibración propuesto. La implementación se llevó a cabo en tres etapas principales, conforme a lo establecido en la metodología del Capítulo 3.

En la primera etapa, se efectuó la calibración intrínseca de los subsistemas que componen la arquitectura híbrida Estéreo–TVS. En particular, se calibró el sistema de estéreo visión mediante el método de Zhang [15], mientras que la calibración del Sistema de Visión Técnica (TVS) requirió el uso de una superficie plana y el sistema de estereovisión para detectar ajustes necesarios tanto mecánicos como computacionales.

Las etapas segunda y tercera correspondieron a la calibración extrínseca entre ambos sistemas. Esta se realizó bajo dos escenarios experimentales distintos: (i) un entorno controlado sobre una mesa óptica, utilizando una superficie plana como elemento auxiliar, y (ii) un entorno móvil, en el que el sistema combinado fue montado sobre un robot móvil para escanear una pared plana mientras se desplazaba y detenía en distintas posiciones de una trayectoria rectilínea. Nuestro método no está diseñado para tomar lecturas en movimiento, por tanto requiere que el robot se detenga para capturar datos.

En esta sección se detallan cada una de estas fases, describiendo los dispositivos utilizados, las condiciones experimentales, la cantidad de experimentos ejecutados y las variaciones introducidas en los factores analizados. Algunos de estos ensayos fueron diseñados siguiendo un enfoque factorial, con el objetivo de analizar la sensibilidad del sistema frente a cambios controlados en parámetros relevantes.

Finalmente, los resultados se presentan en forma de tablas y gráficas, junto con su análisis.

4.1. Calibración intrínseca del sistema Estéreo-TVS

El método de calibración desarrollado en este trabajo fue aplicado sobre dos subsistemas: el Sistema de Visión Técnica (TVS), específicamente su prototipo denominado "Mini TVS", y un sistema de visión estéreo GXIVision con cámaras LSM22100.

Los parámetros intrínsecos del TVS incluyen elementos que requiere tanto ajustes mecánicos como computacionales. Entre los factores con ajustes mecánicos se encuentran la inclinación individual y alineación conjunta de los espejos dispuestos en el posicionador y en la apertura. Estos elementos son críticos para garantizar la correcta triangulación del haz láser. Por otra parte, los factores que requieren ajustes computacionales se centran en la caracterización del comportamiento de los motores a pasos, incluyendo la estimación de la relación grados por paso y la cuantificación del backlash en los cuatro sentidos de movimiento del posicionador.

El prototipo del TVS empleado se muestra en la Figura 4.1. Este cuenta con una separación fija de 20 cm entre el posicionador y la apertura (línea base). El direccionamiento angular del haz es controlado por dos motores a pasos gobernados por un microcontrolador Teensy 4.1, el cual está conectado a una computadora con procesador Intel i5-10300H a 2.50 GHz, con sistema operativo Windows, donde se ejecutan los algoritmos desarrollados en Python para el control del sistema y el cálculo de coordenadas tridimensionales.

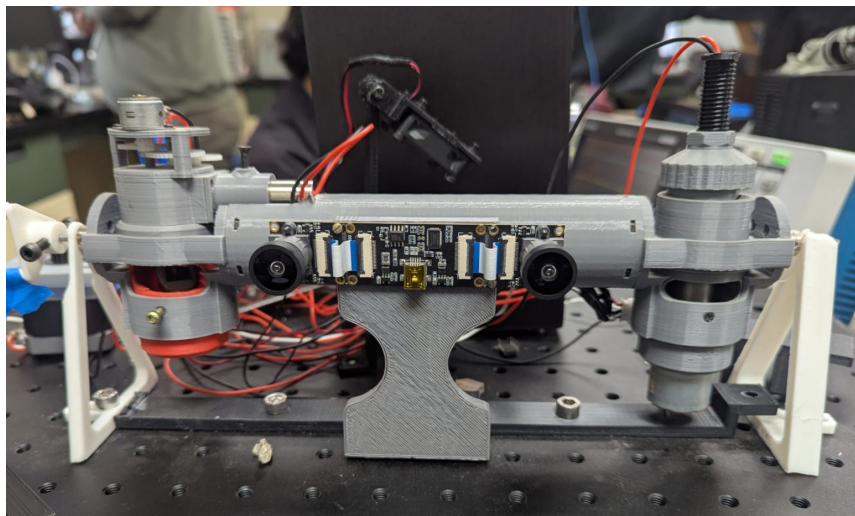


Figura 4.1: Prototipo del Sistema de Visión Técnica (TVS) llamado "Mini TVS".

El software utilizado se encarga de la adquisición de datos y del cálculo de coordenadas tridimensionales, manteniendo sincronía con el Teensy mediante comandos seriales. Esto permite controlar con precisión el posicionamiento horizontal y vertical de la mancha láser emitida por el diodo de 808 nm incorporado en el sistema.

Por su parte, el sistema de estereovisión utilizado está compuesto por un módulo GXIVision (ver Figura 4.2), el cual integra dos cámaras LSM22100 cuyo procedimiento de calibración se describe en la siguiente sección.



Figura 4.2: Módulo de estereovisión GXIVision con dos cámaras LSM22100.

A diferencia de los patrones impresos comúnmente empleados en la calibración de cámaras, el método propuesto para la calibración del sistema híbrido únicamente requiere una superficie plana sobre la cual se proyecta el láser del TVS. Para este propósito, se utilizó como herramienta auxiliar un tablero blanco de papel bond de 40×40 cm. Si bien en la práctica no existen superficies perfectamente planas, se asumió que el tablero proporcionaba un grado de planitud suficiente para aproximarse a un plano ideal. Cabe destacar que este tablero no requiere dimensiones conocidas ni marcas características, lo que diferencia a la metodología presentada de los métodos tradicionales descritos en el Capítulo 1.

Tanto el TVS como el sistema estéreo y el patrón plano fueron montados sobre una mesa óptica de 80×50 cm (Figura 4.3). En este montaje, se definió un sistema de referencia global con origen coincidente con el centro de la línea base del TVS, alineado con su eje longitudinal y concéntrico a su cilindro principal, como se esquematiza en la Figura 4.4.



Figura 4.3: Configuración experimental con el TVS, el sistema estéreo y el patrón plano montados sobre una mesa óptica.

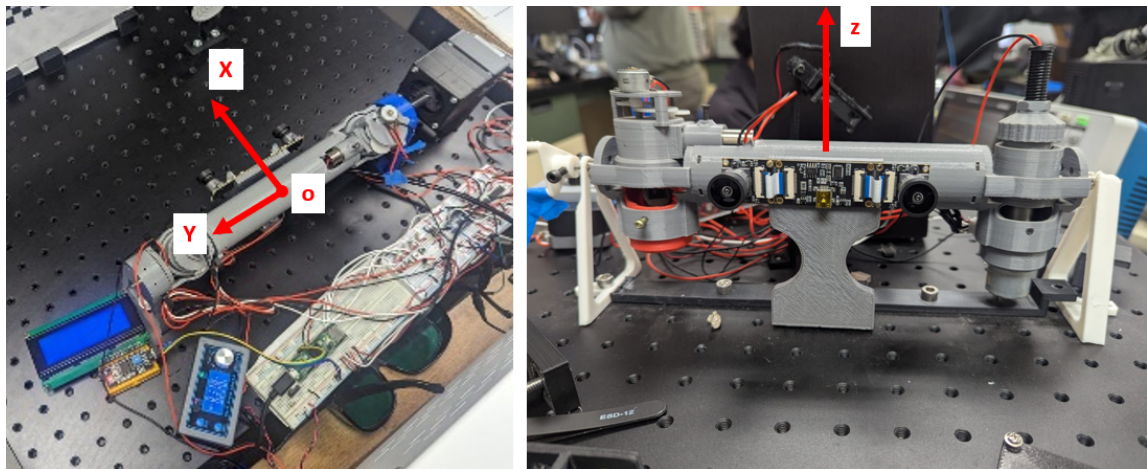


Figura 4.4: Sistema de referencia asociado a la mesa óptica y al TVS, empleado como marco global para la interpretación de resultados.

4.1.1. Calibración del sistema de estereovisión

Como paso preliminar para la calibración intrínseca del Sistema de Visión Técnica (TVS), según la configuración 1 descrita en el Capítulo 3, fue necesario realizar una calibración completa del sistema de visión estéreo utilizando un procedimiento externo. Para ello, se implementó el método de Zhang [15], basado en la observación de un tablero de ajedrez plano con esquinas bien definidas. Se utilizó un tablero de ajedrez con una configuración de 10 columnas \times 7 filas, lo que permite la detección de 54 esquinas internas. En la Figura 4.5 se muestra el tablero empleado para la calibración, que durante la adquisición de imágenes fue desplazado e inclinado en diferentes posiciones, conforme a las recomendaciones establecidas en [66], con el fin de cubrir de manera adecuada el campo visual (FOV) de las cámaras.

Se capturaron un total de 20 poses distribuidas estratégicamente, lo que garantiza una amplia cobertura espacial y gran variabilidad angular suficiente para alimentar el proceso de optimización no lineal. Este proceso permite minimizar el error de reproyección al ajustar simultáneamente los parámetros intrínsecos y la posición relativa entre cámaras.

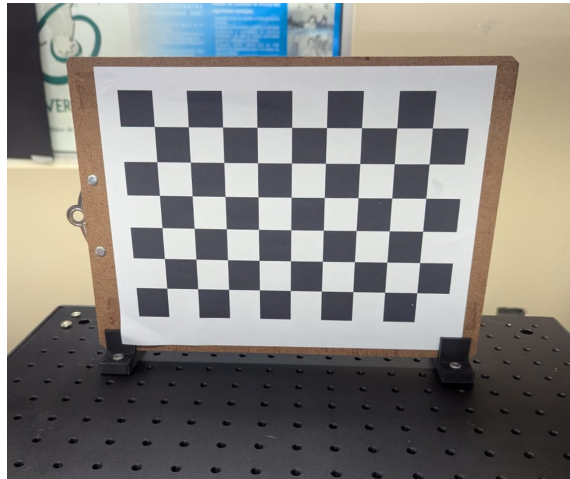


Figura 4.5: Tablero de ajedrez utilizado en la fase inicial del método desarrollado para la calibración del sistema combinado Estéreo-TV S.

Como resultado, se obtuvieron las siguientes matrices intrínsecas de cada cámara, que modelan las propiedades ópticas de sus sistemas de proyección según el modelo pinhole. Estas matrices \mathbf{K}_l y \mathbf{K}_r caracterizan los parámetros internos de la cámara izquierda y derecha, respectivamente:

$$\mathbf{K}_l = \begin{bmatrix} 959.9415958 & 0 & 677.38441068 \\ 0 & 961.75120577 & 365.74125994 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.1)$$

$$\mathbf{K}_r = \begin{bmatrix} 962.55566094 & 0 & 632.91245795 \\ 0 & 964.42494976 & 356.36908189 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.2)$$

Posteriormente, se efectuó la calibración estereoscópica para estimar la relación geométrica entre las dos cámaras. Esto permitió obtener la matriz esencial \mathbf{E} , la cual encapsula la relación epipolar entre los puntos correspondientes en ambas vistas, considerando cámaras previamente calibradas, como se muestra en la Ecuación (4.3):

$$\mathbf{E} = \begin{bmatrix} 3.0122 \times 10^{-3} & -0.3336 & -0.969975 \\ -0.2119 & 4.6115 & 169.9622 \\ 0.9921 & -169.9619 & 4.6145 \end{bmatrix} \quad (4.3)$$

Asimismo, se derivó la matriz fundamental \mathbf{F} , que relaciona puntos homólogos entre imágenes sin requerir conocimiento intrínseco. Su cálculo se realiza a partir de la relación $\mathbf{F} = \mathbf{K}_r^{-T} \mathbf{E} \mathbf{K}_l^{-1}$, y su expresión numérica se presenta en la Ecuación (4.4):

$$\mathbf{F} = \begin{bmatrix} 3.2600 \times 10^{-9} & -3.6036 \times 10^{-7} & -8.7812 \times 10^{-4} \\ -2.2886 \times 10^{-7} & 4.9718 \times 10^{-6} & 0.1746 \\ 1.1131 \times 10^{-3} & -1.7826 & 6.8937 \end{bmatrix} \quad (4.4)$$

A partir de la matriz esencial, se aplicó la descomposición mediante valores singulares (SVD) para

recuperar la matriz de rotación \mathbf{R} y el vector de traslación \mathbf{T} , los cuales describen la transformación rígida entre el sistema de referencia de la cámara izquierda y el de la cámara derecha. Esta transformación se presenta en la Ecuación (4.5):

$$\mathbf{R} = \begin{bmatrix} 0.99993 & 0.01157 & -0.00087 \\ 0.01159 & -0.99956 & 0.02714 \\ -0.00056 & -0.02714 & -0.99963 \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} -0.0850 \\ -0.00049 \\ 0.00015 \end{bmatrix} \text{ m} \quad (4.5)$$

Finalmente, se generó la matriz de reproyección estéreo \mathbf{Q} , que permite transformar las coordenadas de píxeles rectificadas y la disparidad en coordenadas tridimensionales del espacio euclidiano. Su expresión se muestra en la Ecuación (4.6):

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & -659.79541 \\ 0 & 1 & 0 & -361.39081 \\ 0 & 0 & 0 & 963.088078 \\ 0 & 0 & 5.88139447 \times 10^{-3} & 0 \end{bmatrix} \quad (4.6)$$

Con el sistema de estereovisión debidamente calibrado, se procedió a emplearlo como herramienta auxiliar para la estimación de los parámetros intrínsecos del TVS, conforme a la metodología descrita en el Capítulo 3.

4.1.2. Inclinación y alineación de los espejos del posicionador y la apertura

Como primer parámetro evaluado en la calibración intrínseca del TVS, se considera la inclinación individual de los dos espejos que controlan la dirección del haz láser: el espejo del elemento posicionador y el espejo de la apertura.

Para la caracterización del espejo del posicionador se utilizó el láser infrarrojo integrado al sistema. En cambio, para la evaluación del espejo de la apertura fue necesario retirar temporalmente el

fotodiodo receptor y colocar en su lugar un láser verde puntual. Esta modificación experimental permitió observar directamente su proyección sobre la superficie plana auxiliar, facilitando la identificación visual del haz. La disposición del láser verde colocado sobre la apertura se muestra en la Figura 4.6.

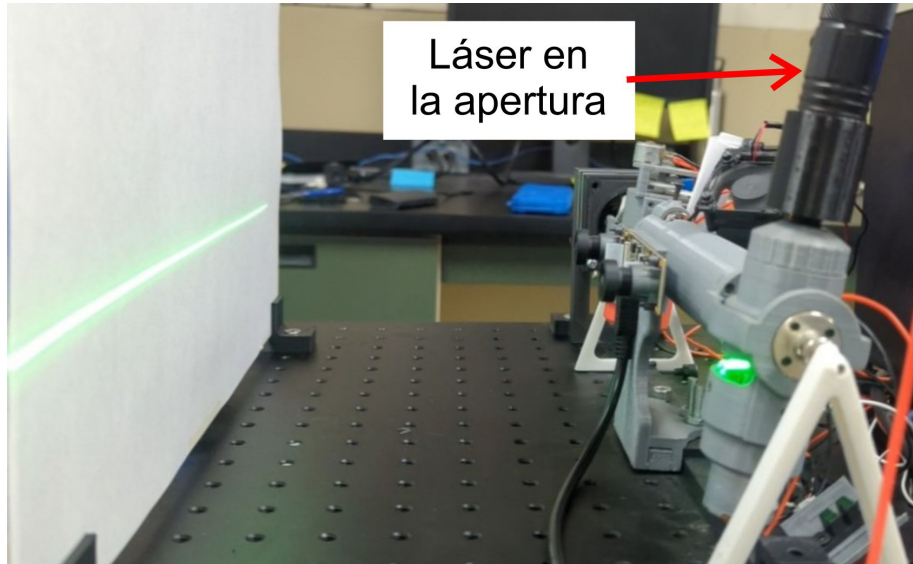


Figura 4.6: Ubicación del láser puntual sobre el espejo de la apertura para su calibración de inclinación.

La evaluación de las desviaciones angulares se realizó empleando el sistema de visión estéreo previamente calibrado, el cual en este caso fue utilizado únicamente como dos cámaras independientes. Cada cámara permitió registrar en coordenadas de imagen (píxeles) la posición proyectada de la mancha láser y del haz verde sobre la superficie plana dispuesta sobre la mesa óptica. De este modo, el análisis se centró en la evolución bidimensional de los centroides de la mancha y líneas detectadas dentro del sistema de coordenadas de cada cámara, sin realizar triangulación ni reconstrucción 3D. La superficie blanca utilizada como plano auxiliar de calibración fue colocada a una distancia fija de 20 cm frente al sistema combinado. En cada experimento se capturaron al menos 20 imágenes, en las cuales el espejo del posicionador fue rotado con el motor a pasos en incrementos de 16 pasos, lo que generó una secuencia de manchas láser distribuidas a lo largo del eje horizontal de la imagen. Este barrido permitió observar el efecto del movimiento angular sobre la posición vertical del haz, siendo este un indicador directo de la inclinación del espejo. El mismo procedimiento se repitió en

24 experimentos.

Durante todas las capturas, el láser verde de la apertura permaneció encendido de manera continua. Este láser proyecta una línea horizontal debido a la rotación continua del espejo de la apertura (30 *RPS*), su presencia en cada imagen se utilizó para medir la inclinación del espejo en la apertura. Cada imagen fue procesada mediante el algoritmo desarrollado en Python que realiza lo mencionado en la metodología del capítulo 3. En la Figura 4.7 se muestra un ejemplo representativo de los resultados obtenidos con la cámara izquierda durante uno de los experimentos. En la imagen se observa la detección simultánea de dos elementos clave: el haz verde proyectado por el láser en la apertura y la mancha láser generada por el láser del posicionador.

El haz verde ha sido segmentado mediante umbralización en el espacio HSV, y sobre los puntos detectados se realizó un ajuste por regresión lineal. La recta obtenida se representa en color rojo. En la esquina superior izquierda se indican el ángulo de inclinación calculado para este experimento (-4.04°) y la ecuación de la recta ajustada en la forma $y = mx + b$, con una pendiente de $m = -0.0707$ y una ordenada al origen de $b = 471.74$.

Por otro lado, la mancha del láser infrarrojo fue identificada mediante segmentación y análisis de brillo. Se calcula su centroide, marcado en la imagen con un punto rojo y sus coordenadas de imagen (805, 496).

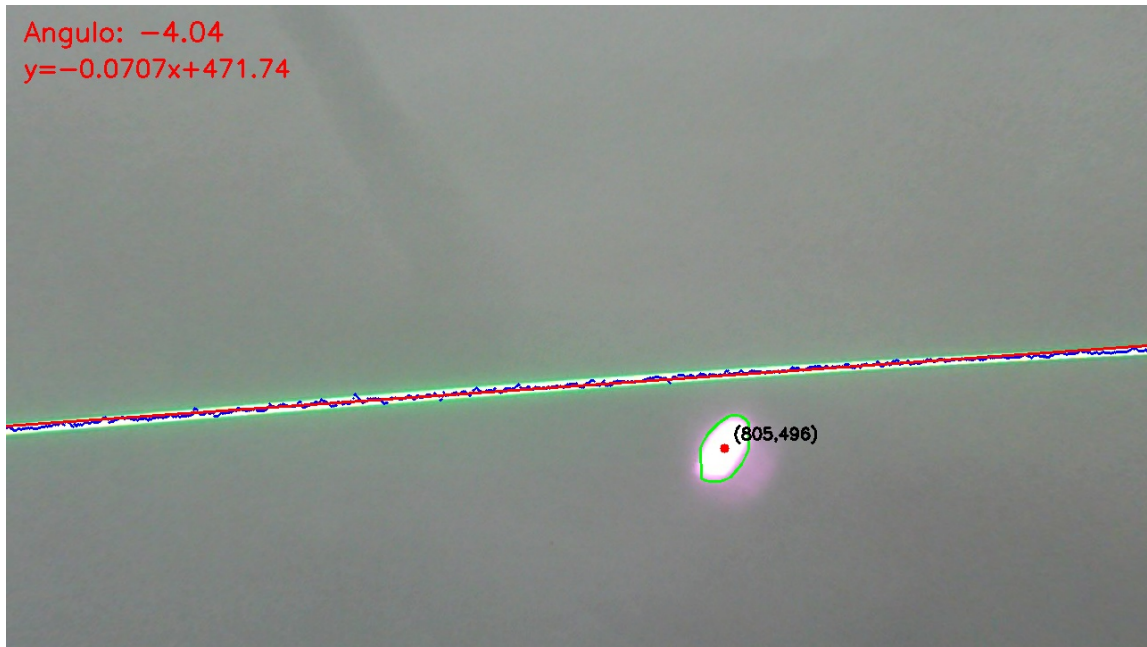


Figura 4.7: Ejemplo de resultado en la evaluación de la inclinación en los espejos del sistema TVS. Se observa el ajuste lineal del haz verde, así como la detección del centroide de la mancha del láser infrarrojo.

Además del análisis individual por imagen, se calculó una recta promedio del haz verde para cada experimento. Esta recta fue determinada como el promedio aritmético de las pendientes m y ordenadas al origen b de las más de 20 líneas ajustadas en las imágenes de su respectivo experimento. En la Figura 4.8 se muestra un experimento representativo, correspondiente a la ejecución número 001. Se grafican los centroides detectados (uno por imagen), junto con la recta de mejor ajuste global (roja) sobre dichos puntos. La línea discontinua naranja representa la recta promedio del haz verde. Este análisis permite evaluar la coherencia angular entre los centroides y el haz proyectado, y detectar posibles desviaciones sistemáticas causadas por la inclinación de los espejos.

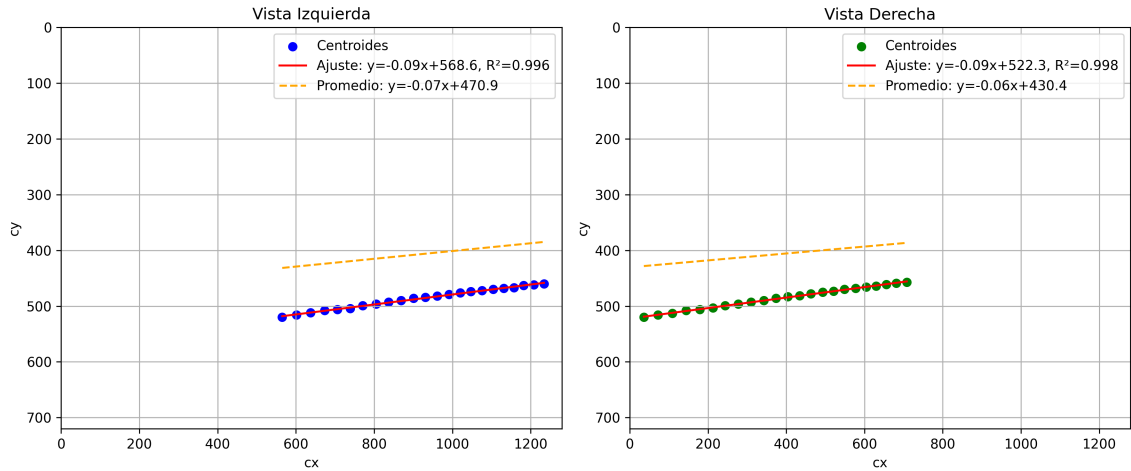


Figura 4.8: Distribución de centroides y líneas ajustadas por imagen en experimento realizado antes del ajuste.

Con base en las 24 ejecuciones experimentales, se calcularon dos métricas de evaluación para cada vista estéreo:

- Δ_m : diferencia entre la pendiente promedio global y la pendiente ajustada por regresión sobre los más de 20 centroides.
- Distancia media entre rectas: medida de consistencia entre la recta promedio y la ajustada globalmente sobre los centroides.

En las Figuras 4.9 y 4.10 se muestran estos resultados antes del ajuste mecánico de los espejos, para las cámaras izquierda y derecha, respectivamente. El análisis puede realizarse con cualquiera de las dos vistas, aunque se incluye en ambas como verificación cruzada.

Métricas de ajuste - Vista Izquierda

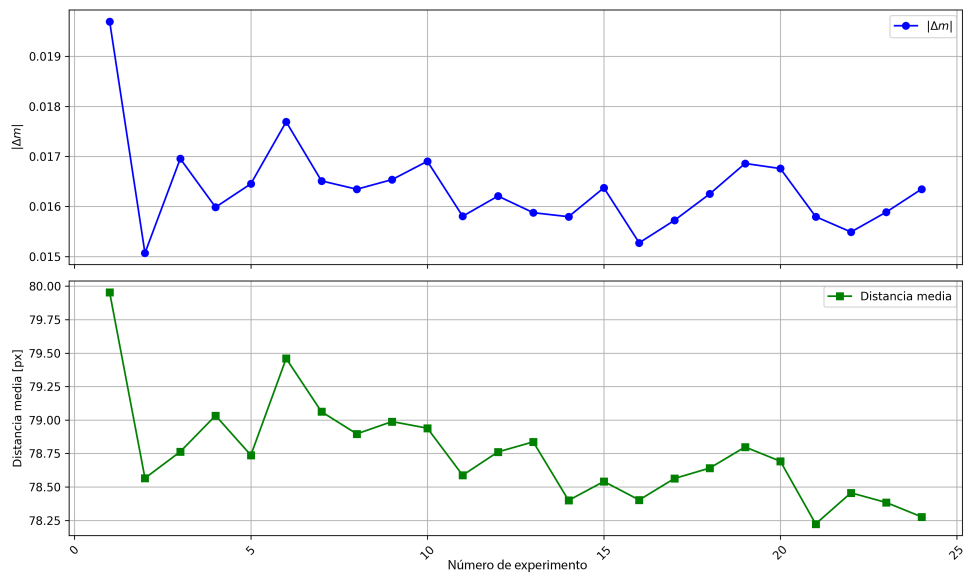


Figura 4.9: Evaluación de inclinación y alineación de espejos previa al ajuste, vista izquierda.

Métricas de ajuste - Vista Derecha

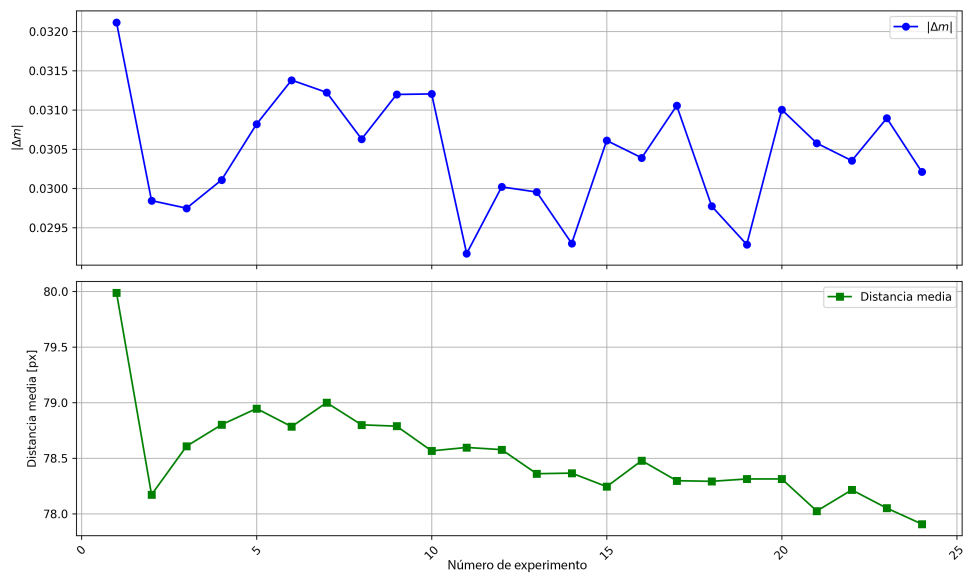


Figura 4.10: Evaluación de inclinación alineación de espejos previa al ajuste, vista derecha.

Luego de estos resultados, se realizó un ajuste mecánico de los espejos, buscando reducir tanto la Δ_m como la distancia promedio. Al realizar el ajuste se intentó lograr tanto una disminución de la inclinación individual como un aumento en la alineación conjunta entre ambos elementos ópticos.

Dicha coalineación se manifiesta cuando las trayectorias proyectadas del haz láser generadas por el espejo del posicionador y el espejo de la apertura resultan coincidentes en el plano.

Desde el punto de vista experimental, esto implica que las pendientes y ordenadas al origen de las rectas ajustadas a los centroides de las manchas láser (producidas por el espejo del posicionador) y al haz verde (generado por el espejo de la apertura) deben ser lo más similares posible.

La Figura 4.11 ilustra un caso representativo posterior al ajuste, en el cual la trayectoria proyectada del haz verde coincide con la ubicación de la mancha generada por el espejo del posicionador. Esta coincidencia visual valida experimentalmente la cercanía a la coalineación entre los dos espejos del sistema TVS, condición necesaria para garantizar la consistencia geométrica en el cálculo posterior de coordenadas tridimensionales.

Por otro lado, la Figura 4.12 muestra el resultado final posterior al ajuste, correspondiente al experimento 001. En ambas vistas se aprecia el aumento de cercanía entre la recta promedio del haz verde y la trayectoria ajustada de los centroides de la mancha láser. Este resultado confirma visual y cuantitativamente la correcta coalineación de los dos espejos del sistema TVS, condición necesaria para el funcionamiento preciso del modelo de triangulación empleado en las siguientes etapas.

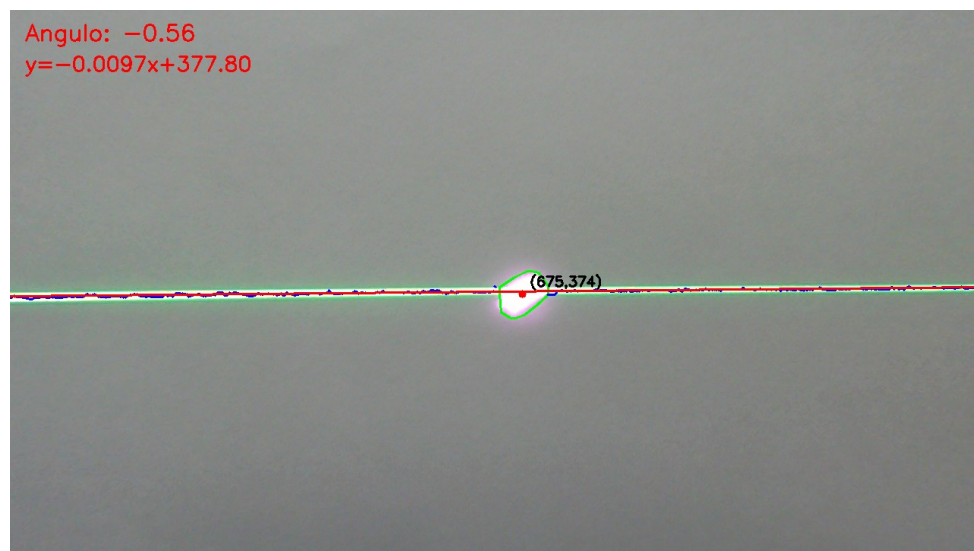


Figura 4.11: Ejemplo representativo de reducción correcta de distancia entre láseres, posterior al ajuste.

Centroides con regresión lineal y recta promedio del haz verde
[001]

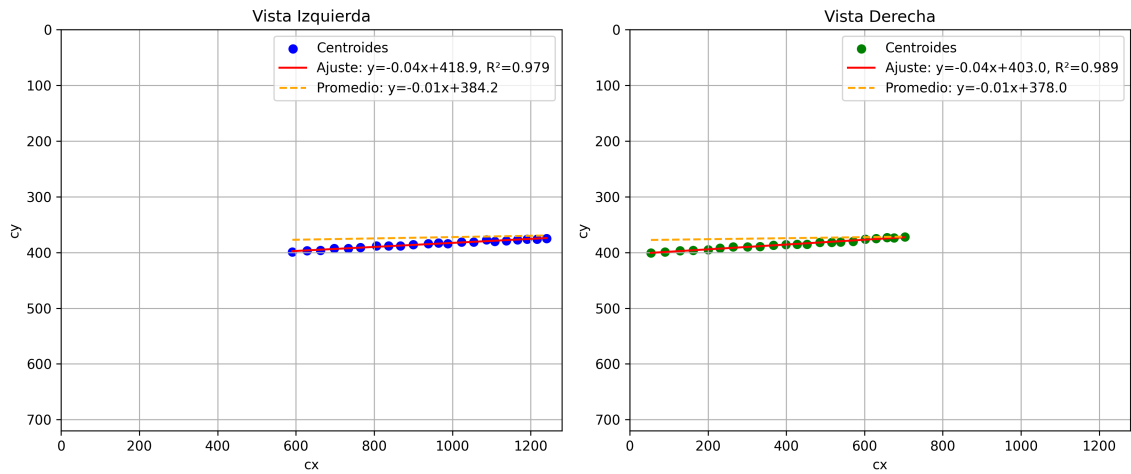


Figura 4.12: Distribución de centroides y líneas ajustadas por imagen en experimento realizado después del ajuste.

Las Figuras 4.13 y 4.14 los resultados de las métricas Δ_m : y Distancia media entre rectas para los 24 experimentos que se realizaron después del ajuste mecánico.



Figura 4.13: Evaluación de inclinación y alineación de espejos después del ajuste, vista izquierda.

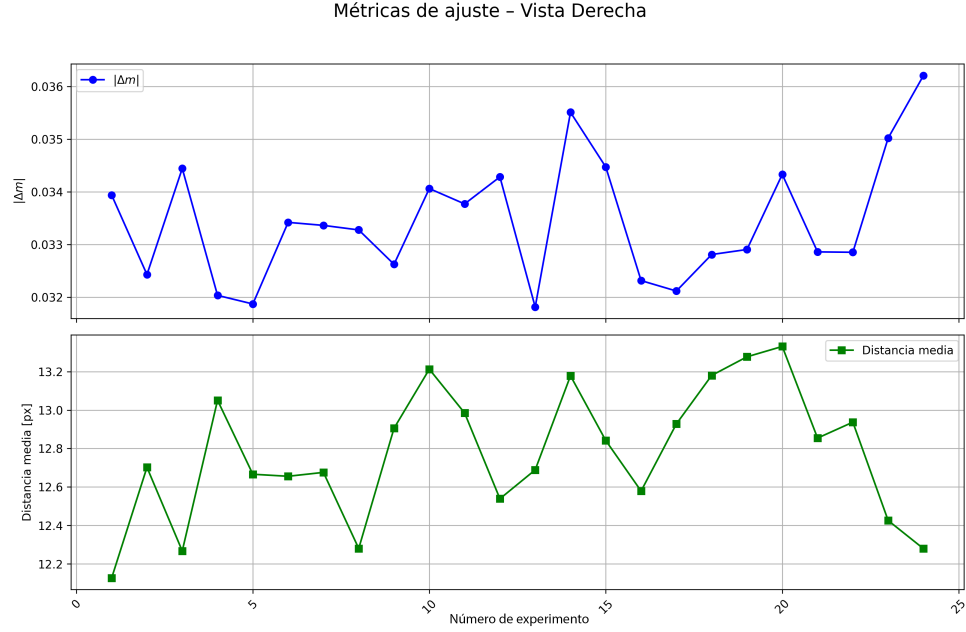


Figura 4.14: Evaluación de inclinación y alineación de espejos después del ajuste, vista derecha.

Para evaluar la coalineación entre espejos, se analizan las métricas de ajuste obtenidas:

- La diferencia entre pendientes, Δ_m , representa la discrepancia angular entre las trayectorias generadas por cada espejo.
- La distancia media entre rectas, expresada en píxeles, cuantifica la separación promedio entre ambas trayectorias proyectadas a lo largo del eje horizontal.

Cabe señalar que la métrica denominada *distancia media entre rectas* no corresponde directamente a la diferencia entre las ordenadas al origen (Δb), sino que representa una estimación integrada de la separación vertical entre la recta promedio del haz verde y la recta ajustada sobre los centroides. Esta métrica se calcula mediante la Ecuación (4.7), en la cual se estima la diferencia vertical punto a punto entre ambas trayectorias.

$$d_i = |(m_{\text{prom}} \cdot x_i + b_{\text{prom}}) - (m_{\text{cent}} \cdot x_i + b_{\text{cent}})|, \quad \text{Distancia media} = \frac{1}{N} \sum_{i=1}^N d_i \quad (4.7)$$

Esto se implementó considerando que al existir una diferencia de pendientes no es pertinente utilizar directamente un valor de (Δb).

Por tanto, esta distancia media actúa como una medida indirecta de la calidad de la coalineación, ya que valores bajos indican que ambas trayectorias —la generada por el espejo del posicionador y la proyectada por el espejo de la apertura— se encuentran próximas y paralelas.

Los resultados presentados en las Figuras 4.13 y 4.14, correspondientes al estado posterior al ajuste, indican una mejora significativa en la reducción de la separación entre rectas, lo cual es un efecto deseado. Sin embargo, logramos apreciar el aumento de Δ_m indicando que a pesar de estar más cerca, los espejos se inclinaron más, uno con respecto al otro, lo cual es un efecto indeseado del ajuste y es atribuido a la poca libertad de ajustes mecánicos que tiene el diseño del prototipo de TVS utilizado.

4.1.3. Caracterización experimental de la relación angular grados/paso

Con el fin de implementar el método de caracterización de la relación grados/paso descrito en la Sección 3, se ejecutaron 24 experimentos independientes para cada una de los cuatro sentidos de movimiento del elemento posicionador del TVS: izquierda, derecha, arriba y abajo. El objetivo fue evaluar experimentalmente la relación angular inducida por cada paso del motor, considerando que se encuentra acoplado a una transmisión mecánica impresa en 3D.

En la Figura 4.15 se muestran los resultados obtenidos al medir las coordenadas tridimensionales en un experimento con desplazamiento hacia la izquierda. Los puntos P_1 y P_2 corresponden a la posición inicial del haz antes del movimiento del motor, mientras que P_3 y P_4 representan la posición del haz después de haber desplazado el motor 200 pasos. Cabe destacar que los puntos P_2 y P_3 se registran en una segunda posición de la superficie plana, que se traslada desde una profundidad inicial aproximada $d_1 = 17.5$ cm hasta $d_2 = 27.5$ cm. El método es robusto ante imprecisiones en la colocación de la superficie plana en la mesa óptica, ya que el análisis se realiza mediante triangulación estéreo con coordenadas del mundo real que también miden la profundidad a la que se encuentra el tablero.

Visualización 3D de P1 a P4 con líneas proyectadas

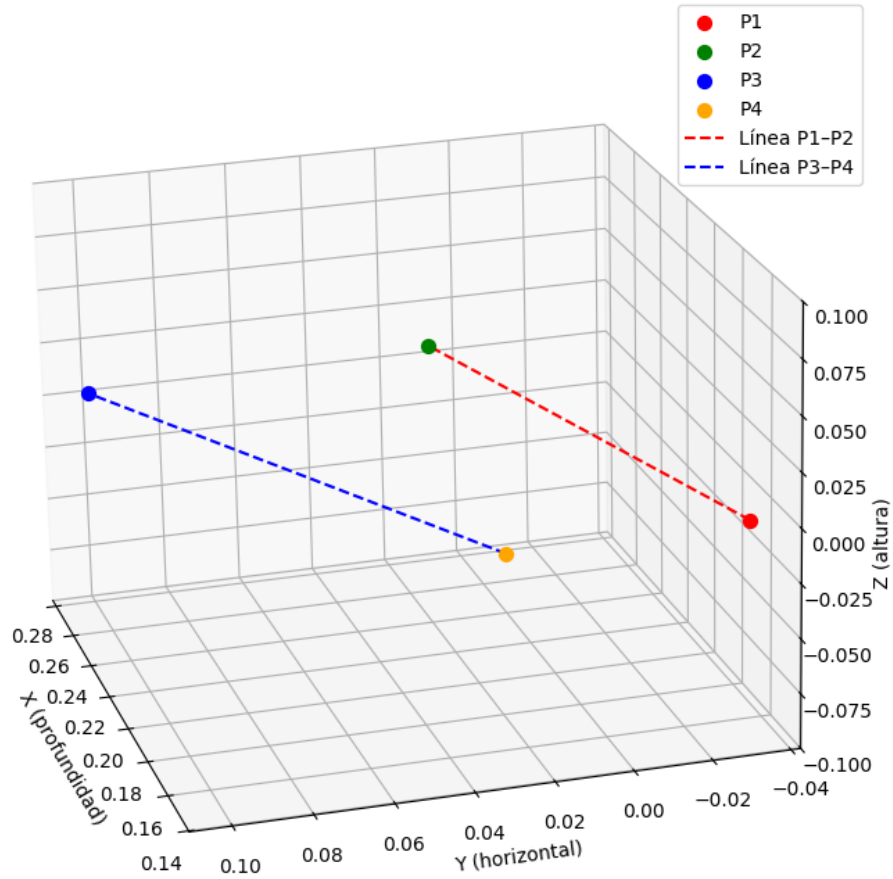


Figura 4.15: Visualización tridimensional de los puntos triangulados en un experimento de desplazamiento hacia la izquierda.

La proyección de estos vectores sobre el plano XY permite calcular la inclinación angular de cada trayectoria utilizando la razón entre el desplazamiento horizontal y el cambio de profundidad, como se muestra en la Ecuación (4.8):

$$\theta_i = \tan^{-1} \left(\frac{Y_2 - Y_1}{X_2 - X_1} \right), \quad \theta_f = \tan^{-1} \left(\frac{Y_3 - Y_4}{X_3 - X_4} \right) \quad (4.8)$$

La Figura 4.16 muestra los triángulos proyectados para el cálculo de los ángulos inicial y final para el mismo experimento analizado.

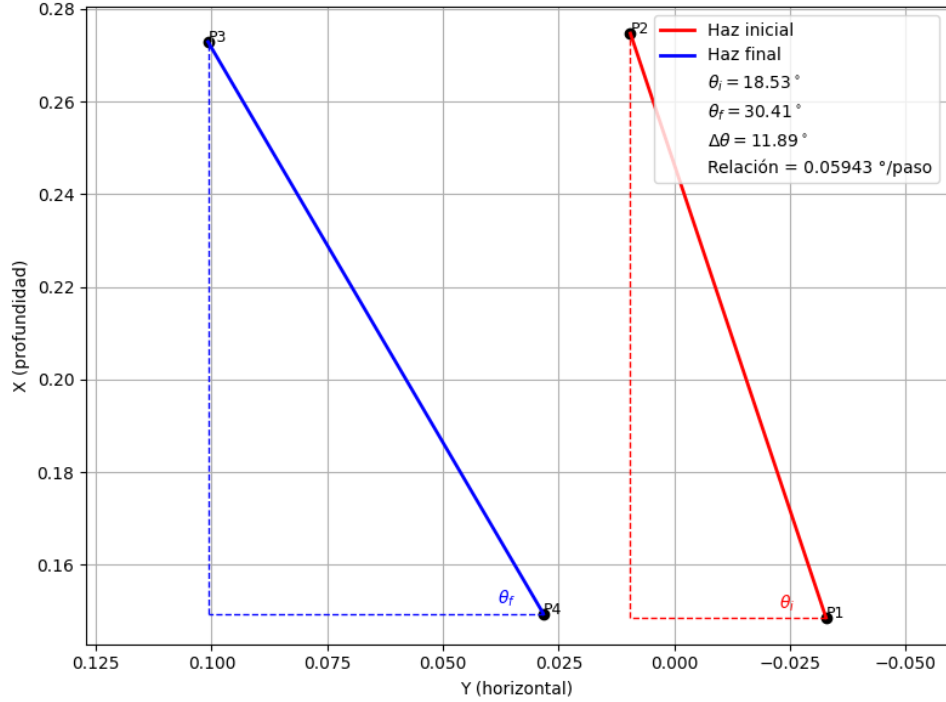


Figura 4.16: Cálculo angular en el plano XY para un experimento individual.

En el ejemplo mostrado, se obtuvo un ángulo inicial de $\theta_i = 18.53^\circ$ y un ángulo final de $\theta_f = 30.41^\circ$, resultando en una diferencia angular de $\Delta\theta = 11.89^\circ$. Dado que en este experimento se aplicaron 200 pasos al motor, la relación estimada se presenta en la Ecuación (4.9):

$$\frac{\Delta\theta}{N_{\text{pasos}}} = \frac{11.89^\circ}{200} = 0.0594^\circ/\text{paso} \quad (4.9)$$

El procedimiento descrito se repitió para las otras tres direcciones de movimiento. En el caso de desplazamientos hacia la derecha, se utilizó el mismo análisis en el plano XY , mientras que para los desplazamientos verticales (hacia arriba y hacia abajo), el análisis se realizó en el plano XZ , conservando la profundidad como eje adyacente y utilizando la componente vertical Z como cateto opuesto en la estimación angular. Esta se calcula mediante la Ecuación (4.10):

$$\theta = \tan^{-1} \left(\frac{Z_{\text{final}} - Z_{\text{inicial}}}{X_{\text{final}} - X_{\text{inicial}}} \right) \quad (4.10)$$

Distribución de resultados de los 24 experimentos en sentido de movimiento

Para evaluar la consistencia de los resultados obtenidos, se representó gráficamente la distribución de la relación grados/paso en cada uno de los sentidos de movimiento. Las Figuras 4.17 a 4.20 muestran las gráficas correspondientes a los 24 experimentos por sentido, incluyendo la media general, así como las bandas de tolerancia asociadas a $\pm 1\sigma$, $\pm 2\sigma$ y $\pm 3\sigma$.

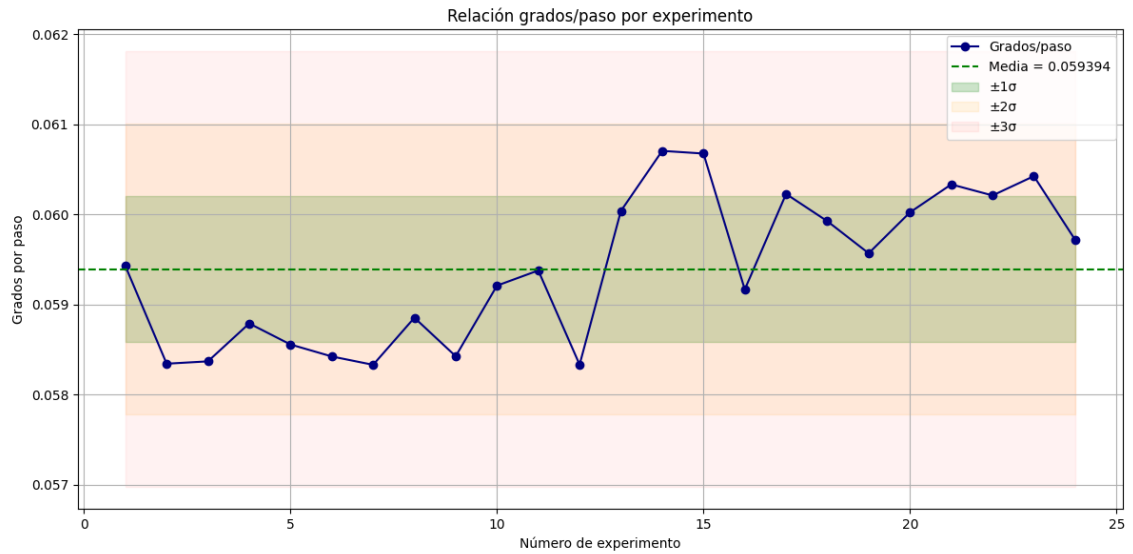


Figura 4.17: Distribución de la relación grados/paso en la dirección izquierda.

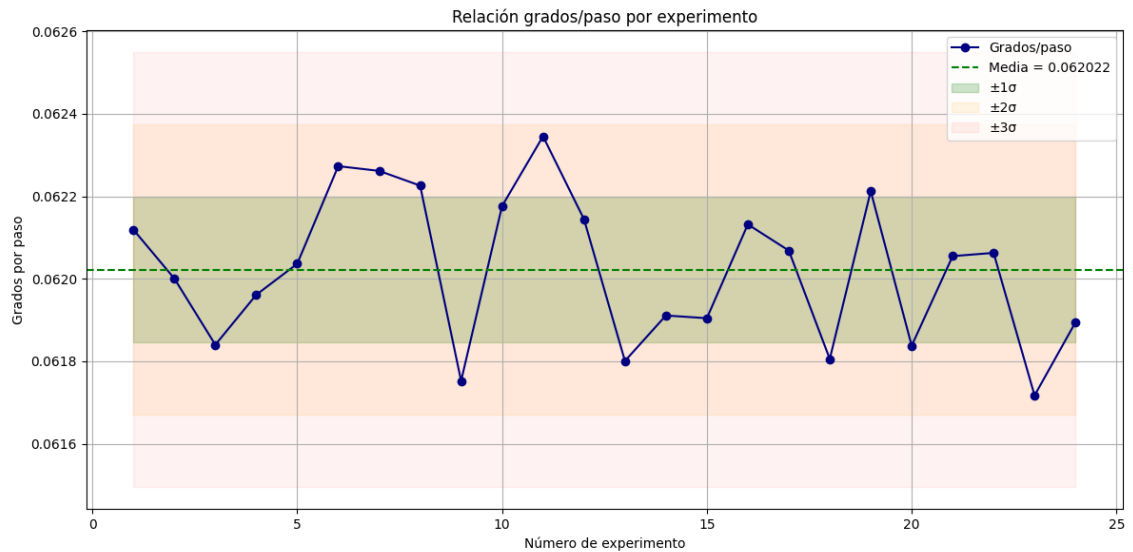


Figura 4.18: Distribución de la relación grados/paso en la dirección derecha.

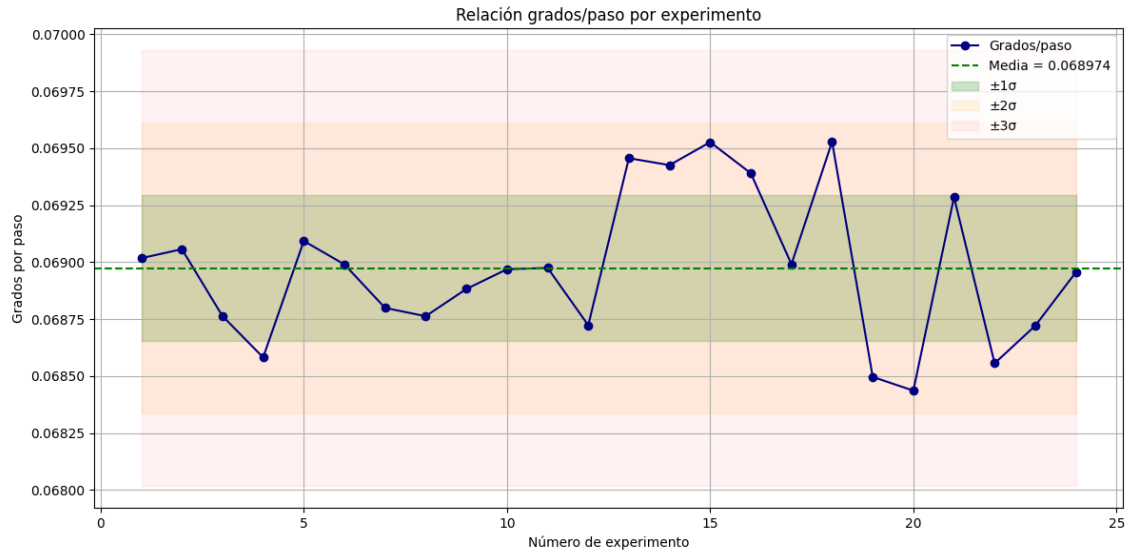


Figura 4.19: Distribución de la relación grados/paso en la dirección arriba.

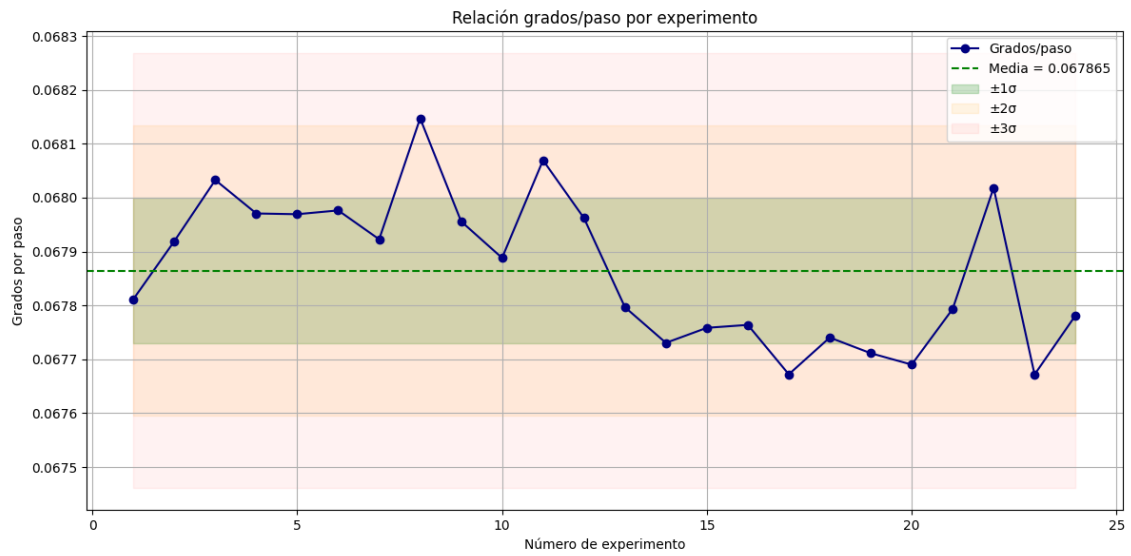


Figura 4.20: Distribución de la relación grados/paso en la dirección abajo.

La Tabla 4.1 presenta un resumen estadístico de los 24 experimentos realizados para cada sentido de movimiento. En ella se reportan la media y la desviación estándar de la relación grados/paso, lo que permite describir cuantitativamente el comportamiento del sistema en cada caso.

Tabla 4.1: Resumen estadístico de la relación grados/paso por dirección.

Dirección	Media [°/paso]	Desviación estándar
Izquierda	0.059394	0.000808
Derecha	0.062023	0.000176
Arriba	0.068974	0.000319
Abajo	0.067865	0.000173

A partir de los resultados obtenidos, se observa que la relación angular grados/paso presenta una alta repetibilidad dentro de cada dirección, como lo evidencia la baja dispersión en torno a la media ($\sigma < 0.001^\circ/\text{paso}$ en todos los casos). Sin embargo, se identifican ligeras asimetrías entre sentidos opuestos del mismo eje, particularmente en el eje horizontal, donde la dirección derecha muestra una media levemente superior ($0.0620^\circ/\text{paso}$) en comparación con la dirección izquierda ($0.0594^\circ/\text{paso}$). Este comportamiento puede atribuirse a efectos sistemáticos introducidos por la transmisión mecánica. En el eje vertical se observa un patrón similar, aunque con valores medios ligeramente más elevados, lo que sugiere que la geometría del mecanismo de posicionamiento vertical podría inducir una relación angular más pronunciada.

Independientemente de la fuente de estos errores, los valores medios estimados pueden incorporarse directamente como factores de corrección en el algoritmo que calcula las coordenadas tridimensionales. De este modo, se logra una compensación computacional eficiente que no requiere modificaciones físicas sobre los componentes mecánicos del sistema, mejorando así la precisión del cálculo de coordenadas sin intervenir en la estructura del hardware.

En la siguiente sección se describe el procedimiento experimental para el método de calibración extrínseca propuesto para el sistema combinado Estéreo-TVS.

4.2. Calibración extrínseca del sistema Estéreo-TVS montado en mesa óptica

Una vez determinados los parámetros intrínsecos del TVS auxiliándonos del sistema de estereovisión, se procedió al cálculo de los parámetros extrínsecos, los cuales definen la relación geométrica entre ambos sistemas y permiten su alineación a un marco de referencia común. En este trabajo, se ejemplifica el procedimiento para transformar la nube de puntos del sistema de estereovisión al sistema de referencia del TVS. Sin embargo, el método desarrollado también admite la conversión inversa, en función de los requerimientos de experimentos futuros.

El método propuesto ha sido diseñado considerando las características específicas de ambos sistemas. De manera análoga a ciertas etapas de la calibración intrínseca descrita en la sección anterior, en este caso se emplea la proyección de la mancha láser emitida por el TVS sobre una superficie plana. A partir de esta proyección, se construye una nube de puntos tridimensionales correspondiente al sistema de estereovisión, cuya alineación con la nube de puntos adquirida directamente con el TVS se determina mediante el ajuste de planos. Este enfoque presenta la ventaja de permitir la detección de la mancha láser incluso en condiciones de oscuridad, lo que lo diferencia de otros métodos descritos en el estado del arte, los cuales requieren condiciones de iluminación controlada para garantizar la correcta captura de imágenes por parte de las cámaras.

A continuación, se detallan los experimentos realizados relacionados con la calibración extrínseca, abarcando desde la descripción de los equipos utilizados y la preparación del escenario experimental hasta el proceso de adquisición de datos y los resultados obtenidos en diversas configuraciones experimentales.

4.2.1. Configuración experimental para la calibración extrínseca

Para la calibración extrínseca, se empleó el prototipo Mini TVS junto con el módulo de estereovisión GXIVision, equipado con dos cámaras LSM22100. Ambos dispositivos fueron montados nuevamente sobre la mesa óptica de 80×50 cm, en conjunto con una superficie plana de 40×40 cm. La disposición

de estos equipos se ilustra en la Figura 4.21.



Figura 4.21: Disposición del equipo en la mesa óptica para la calibración extrínseca.

Para evaluar la eficacia del método en distintas condiciones experimentales, se variaron diversos parámetros, incluyendo la posición del patrón en la mesa óptica, como se muestra en la Figura 4.22. En cada una de las tres ubicaciones del patrón, se realizaron ocho experimentos distintos, repitiendo cada combinación dos veces con el fin de garantizar la fiabilidad de los resultados. Por otro lado, la Tabla 4.2 resume los factores y niveles considerados en el diseño experimental. En total, se llevaron a cabo 48 experimentos.

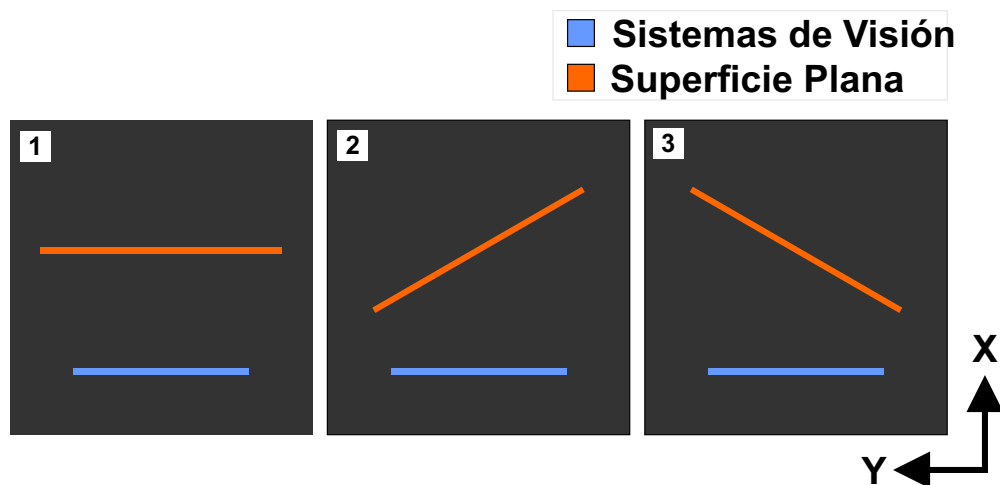


Figura 4.22: Ubicaciones del patrón en la mesa óptica durante la experimentación.

Tabla 4.2: Factores y niveles empleados en diseño experimental para la evaluación del método de calibración extrínseca sobre mesa óptica.

Factor	Niveles
Iluminación	Luz ambiental (634 lx), Oscuridad (5 lx)
Número de puntos (N)	4, 8, 16, 32

Los factores seleccionados en el diseño experimental responden a criterios prácticos orientados a evaluar la robustez y eficiencia del método propuesto. El parámetro de iluminación permite validar la capacidad del sistema para operar sin apoyo de fuentes luminosas externas, aprovechando la alta intensidad luminosa de la mancha láser, lo cual resulta especialmente útil en entornos confinados y oscuros, como tuberías o ductos de ventilación. Por otro lado, los niveles del número de puntos fueron definidos para analizar el impacto de la densidad de muestreo sobre la estimación extrínseca. A diferencia de métodos que requieren grandes volúmenes de datos para garantizar precisión, en este trabajo se busca demostrar que es posible obtener resultados confiables con cantidades reducidas de puntos, lo cual disminuye significativamente el tiempo de adquisición y procesamiento sin comprometer la fidelidad geométrica del ajuste.

4.2.2. Adquisición de datos y construcción de nubes de puntos

Durante cada experimento, el tablero plano se posicionó dentro del campo de visión (FOV) de ambos sistemas, lo cual es un requisito fundamental para el método propuesto. Durante el proceso de adquisición de datos, el láser del TVS se proyecta sobre la superficie plana, realizando un escaneo con un número predefinido de puntos. En cada posición del láser, el TVS registra las coordenadas tridimensionales del punto correspondiente y las almacena, generando así su respectiva nube de puntos, denotada como P_{cTVS} .

Simultáneamente, en cada posición del láser, el sistema de estereovisión captura imágenes de la mancha del haz con ambas cámaras. Posteriormente, se aplican técnicas de rectificación para corregir las imágenes y se segmenta la mancha láser mediante umbralización para identificar su contorno. A partir de esta segmentación, se calcula el centroide de la mancha en coordenadas de píxeles. Finalmente, a través de un proceso de triangulación estereoscópica, estas coordenadas se transforman

en un sistema tridimensional, obteniendo la nube de puntos del sistema de estereovisión, denotada como P_{cS} .

La segmentación e identificación del centroide de la mancha láser se ilustra en la Figura 4.23, tanto en condiciones de luz ambiental como en total oscuridad. Esto demuestra la efectividad del método para la detección del láser sin depender de patrones estructurados en la superficie de calibración, lo que permite una adquisición robusta de los puntos tridimensionales en diversas condiciones de iluminación.

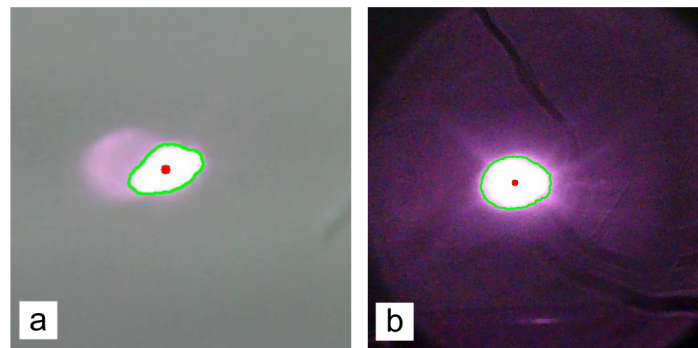


Figura 4.23: Detección del contorno y centroide de la mancha láser bajo diferentes condiciones de iluminación. a): luz ambiental, b): oscuridad.

Sabemos que las nubes P_{cTVS} y P_{cS} están compuestas por puntos homólogos, sin embargo cada sistema los captura desde su propio marco de referencia. Un ejemplo de las nubes de puntos adquiridas durante la experimentación se presenta en la Figura 4.24. Se logra observar que la nube de puntos obtenida por el sistema de estereovisión presenta un desplazamiento y una rotación con respecto a la nube de puntos del TVS, lo que justifica la necesidad del proceso de calibración extrínseca.

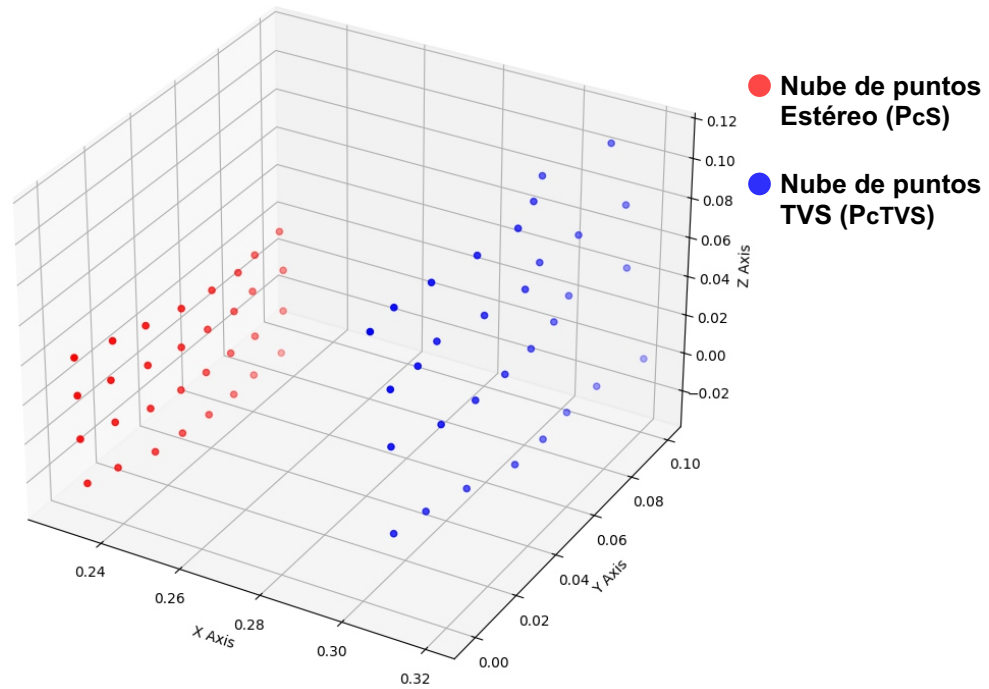


Figura 4.24: Ejemplo de nubes de puntos obtenidas en la experimentación.

4.2.3. Aplicación del método de transformación

La Fig. 4.25 se obtuvo a partir de un experimento y presenta las nubes de puntos junto con sus respectivos planos ajustados y los vectores ortonormales que sirven como sistemas de referencia asociados a cada plano. La nube de puntos del sistema de estereovisión fue rotada utilizando la Ecuación 3.25 y trasladada mediante la Ecuación 3.29.

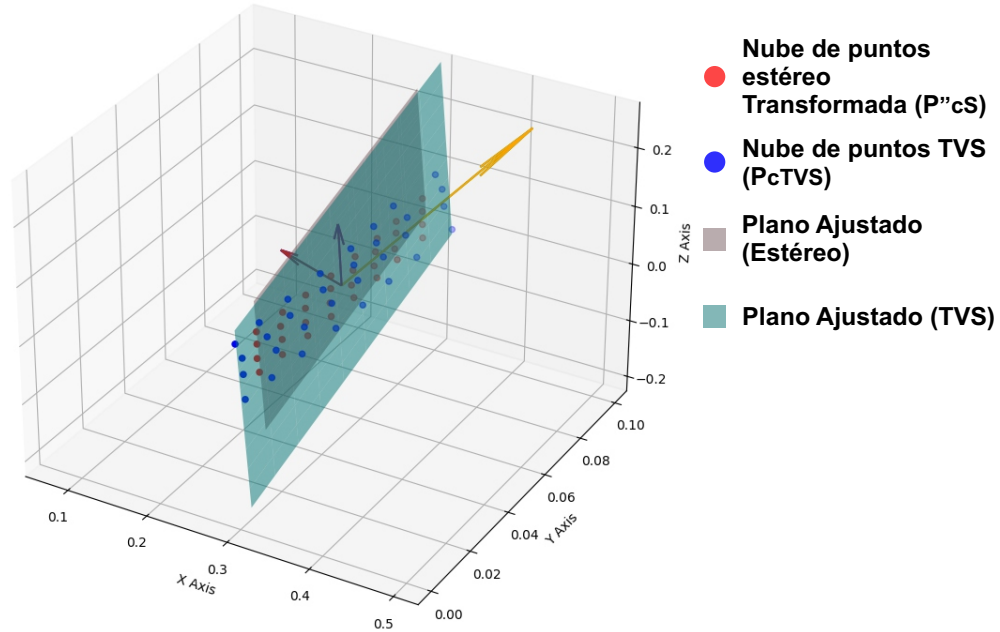


Figura 4.25: Nube de puntos original del TVS y nube de puntos transformada del sistema de estereovisión con planos ajustados y vectores de referencia.

4.2.4. Error de correspondencia posterior a la transformación

Se analizó la correspondencia entre las nubes de puntos con el propósito de evaluar la precisión y robustez del método de calibración diseñado en este trabajo. En un escenario ideal, ambas nubes deberían alinearse perfectamente tras la transformación, esto debido a que ambas nubes están constituidas completamente por puntos homólogos; sin embargo, se identificaron discrepancias. Dichas diferencias fueron utilizadas para calcular el error medio en los ejes (x, y, z) , obteniendo el promedio de las desviaciones entre los puntos transformados del sistema de estereovisión y sus homólogos en la nube original del TVS.

El error de correspondencia se cuantifica como el promedio de las diferencias absolutas entre los puntos de la nube transformada de estereovisión y sus correspondientes en la nube del TVS, tal como se muestra en la Ecuación (4.11):

$$E_{avg} = \frac{1}{N} \sum_{i=1}^N |P_{TVS,i} - P_{S,i}^{transformado}| \quad (4.11)$$

donde E_{avg} representa el error promedio, N es el número total de puntos, $P_{TVS,i}$ denota las coordenadas del punto i en la nube del TVS, y $P_{S,i}^{transformado}$ corresponde a las coordenadas del punto equivalente en la nube del sistema de estereovisión, una vez transformado mediante la matriz de rotación y el vector de traslación que lo alinean con el sistema de referencia del TVS.

Los resultados obtenidos de los 48 experimentos realizados se presentan en las Figuras 4.26, 4.27 y 4.28, las cuales muestran el error registrado en cada eje. En estas gráficas se comparan directamente las dos condiciones de iluminación evaluadas: luz ambiental (línea azul) y oscuridad (línea roja). Cabe señalar que cada condición contempla 24 experimentos independientes, de modo que el eje horizontal se limita a 24 posiciones para facilitar la comparación directa, mientras que el fondo segmentado en tres bloques de ocho experimentos representa las distintas ubicaciones de la superficie plana sobre la mesa óptica. Esta visualización permite contrastar el desempeño del método bajo diferentes condiciones de iluminación, evidenciando su efectividad incluso en entornos con visibilidad reducida. Adicionalmente, la Tabla 4.3 resume el error promedio y la desviación estándar obtenidos en los 48 experimentos para cada uno de los ejes, siendo el eje X el correspondiente a la medición de profundidad.

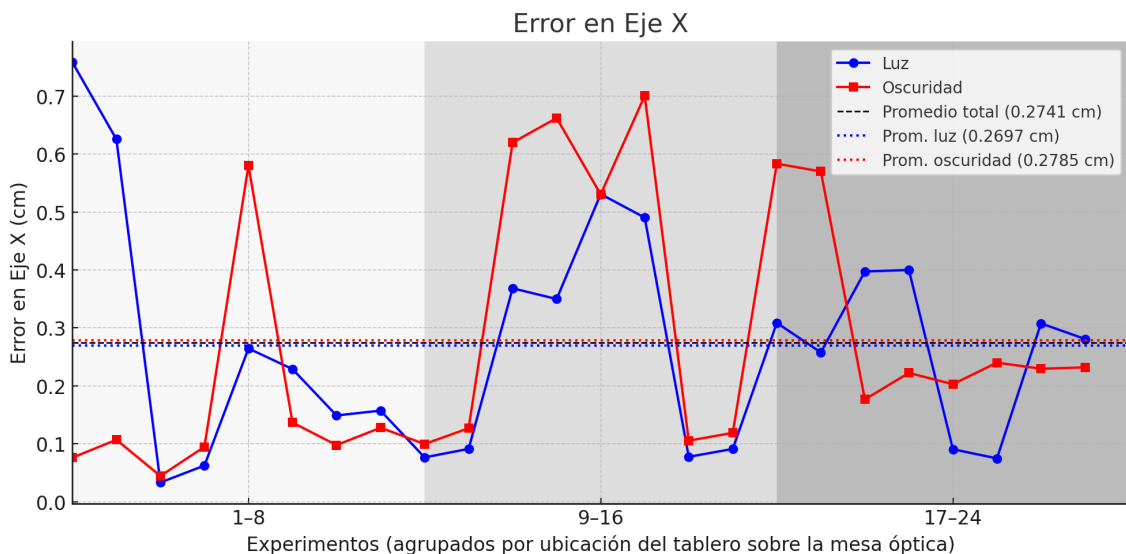


Figura 4.26: Error de correspondencia en el eje X en experimentos sobre mesa óptica.

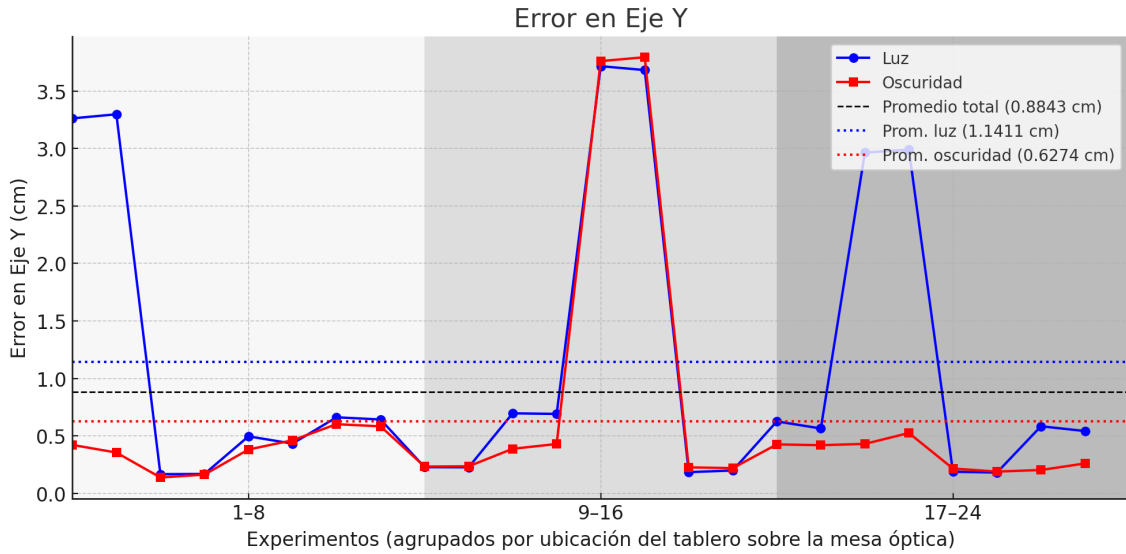


Figura 4.27: Error de correspondencia en el eje Y en experimentos sobre mesa óptica.

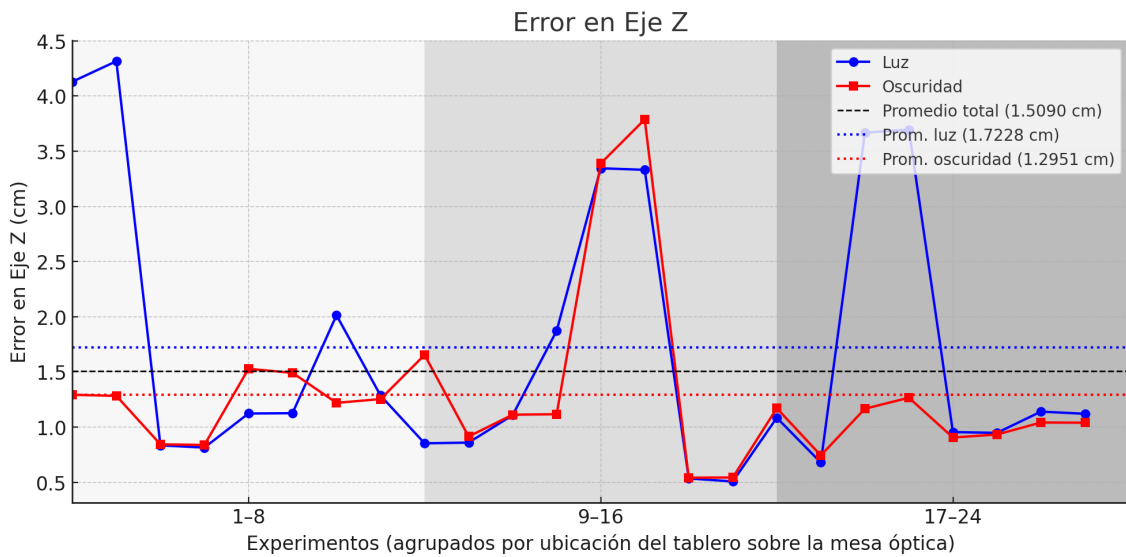


Figura 4.28: Error de correspondencia en el eje Z en experimentos sobre mesa óptica.

Tabla 4.3: Resumen de resultados de calibración extrínseca considerando 48 experimentos sobre mesa óptica.

Eje	Media (cm)	Desviación estándar (cm)
X	0.2741 cm	0.2072 cm
Y	0.8843 cm	1.1771 cm
Z	1.5090 cm	1.0490 cm

Al analizar los resultados separando los niveles de iluminación, se observó que tanto en condiciones

de luz ambiental como en oscuridad, los errores promedio y su variabilidad se mantuvieron bajos. Bajo condiciones de luz ambiental, los errores promedio fueron de 0.2697 cm , 1.1411 cm y 1.7228 cm en los ejes X , Y y Z , respectivamente, con desviaciones estándar de 0.1946 cm , 1.306 cm y 1.254 cm . Mientras tanto, en condiciones de oscuridad, los errores promedio fueron de 0.2785 cm , 0.6274 cm y 1.2951 cm , con desviaciones estándar de 0.2231 cm , 0.979 cm y 0.761 cm , respectivamente.

De manera similar, al analizar los resultados en función del número de puntos, los errores promedio y sus desviaciones estándar fueron los siguientes: para 4 puntos, 0.2510 cm con una desviación estándar de 0.2561 cm en el eje X , 0.812 cm con 1.160 cm en el eje Y y 1.635 cm con 1.235 cm en el eje Z ; para 8 puntos, 0.2438 cm , 0.468 cm y 1.272 cm con desviaciones estándar de 0.2090 cm , 0.1986 cm y 0.7615 cm , respectivamente; para 16 puntos, 0.3637 cm , 1.484 cm y 1.636 cm con desviaciones estándar de 0.2283 cm , 1.160 cm y 1.205 cm ; y para 32 puntos, 0.2379 cm , 0.8306 cm y 1.490 cm , con desviaciones estándar de 0.1826 cm , 0.5629 cm y 1.029 cm , respectivamente.

Finalmente, al evaluar las posiciones experimentales, se obtuvieron errores promedio bajos y consistentes en todas las configuraciones. En la Posición 1, los errores promedio fueron de 0.3467 cm , 1.980 cm y 2.527 cm en los ejes X , Y y Z , respectivamente, con desviaciones estándar de 0.2348 cm , 1.521 cm y 1.257 cm . En la Posición 2, los errores promedio fueron de 0.1018 cm , 0.1977 cm y 0.8429 cm , con desviaciones estándar de 0.0530 cm , 0.0296 cm y 0.2693 cm , respectivamente. Finalmente, en la Posición 3, los errores promedio fueron de 0.3738 cm , 0.4750 cm y 1.1565 cm , con desviaciones estándar de 0.1694 cm , 0.1388 cm y 0.2833 cm .

Los resultados obtenidos permiten concluir que el método propuesto es robusto frente a condiciones adversas de operación. A diferencia de enfoques que dependen de iluminación controlada o patrones estructurados, el uso del haz láser como referencia permite lograr una alineación precisa incluso en completa oscuridad. Asimismo, se observó que el método mantiene un rendimiento aceptable aun con un número reducido de puntos de escaneo, lo que resulta útil para escenarios en los que el tiempo de adquisición debe minimizarse. Finalmente, la baja dispersión de errores entre las distintas posiciones del patrón confirma que la técnica es consistente y confiable bajo variaciones espaciales, lo cual refuerza su aplicabilidad en contextos de operación dinámica o en entornos no controlados.

4.3. Calibración extrínseca del sistema Estéreo-TVS montado en robot con movimiento intermitente

Una vez validada la efectividad del procedimiento de calibración extrínseca sobre el montaje en mesa óptica, se procedió a extender la evaluación al entorno operativo de un robot móvil. En esta etapa, la calibración extrínseca se llevó a cabo en un escenario de desplazamiento intermitente, en el cual el robot se detiene en intervalos específicos para realizar las mediciones.

La Figura 4.29 muestra el robot móvil utilizado en la fase experimental, junto con los dispositivos asociados para la adquisición de datos y el monitoreo de señales y coordenadas. La gestión del sistema TVS se mantuvo mediante una placa de desarrollo Teensy 4.1, que en esta configuración se comunica a través de una interfaz serial con un Arduino Uno, el cual controla el desplazamiento del robot. El sistema estereoscópico sigue estando basado en el módulo GXIVision con dos cámaras LSM22100.

El procesamiento y visualización de los datos se llevaron a cabo en una computadora con un procesador Intel Core i5-10300H @ 2.50 GHz, la cual ejecuta el software de control encargado de sincronizar la captura de imágenes del TVS y almacenar simultáneamente las nubes de puntos generadas por ambos sistemas.

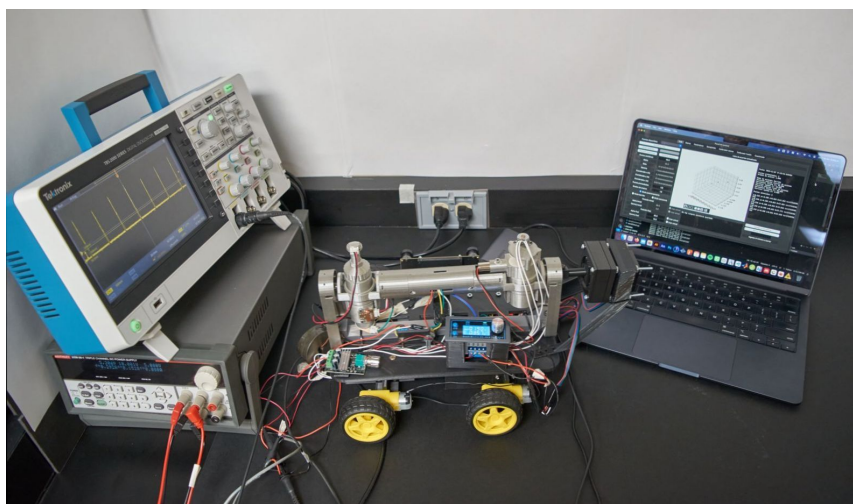


Figura 4.29: Configuración experimental: robot móvil con sistema combinado Estéreo-TVS, junto con la computadora utilizada para el control del robot y el osciloscopio Tektronix TBS2000 para el monitoreo de la señal.

En esta fase, se realizó un diseño experimental considerando la variación de tres factores clave. La Tabla 4.4 presenta estos factores junto con sus respectivos niveles de variación. Cada combinación se repitió tres veces para garantizar la fiabilidad de los resultados, dando un total de 72 experimentos.

Factor	Niveles
Iluminación	Luz ambiental (354 lx), Oscuridad (2 lx)
Distancia	Corta (20 cm), Media (30 cm), Larga (40 cm)
Posiciones (k)	3, 4, 5, 6

Tabla 4.4: Factores y niveles empleados en diseño experimental para la evaluación del método de calibración extrínseca sobre robot con movimiento intermitente.

Las Figuras 4.30, 4.31 y 4.32 muestran los errores registrados en los ejes X , Y y Z , respectivamente, para las distintas combinaciones de distancia e iluminación. Siendo el eje X el correspondiente a la profundidad. El eje horizontal se limita a 36 experimentos, correspondientes a cada línea de tendencia (72 en total).

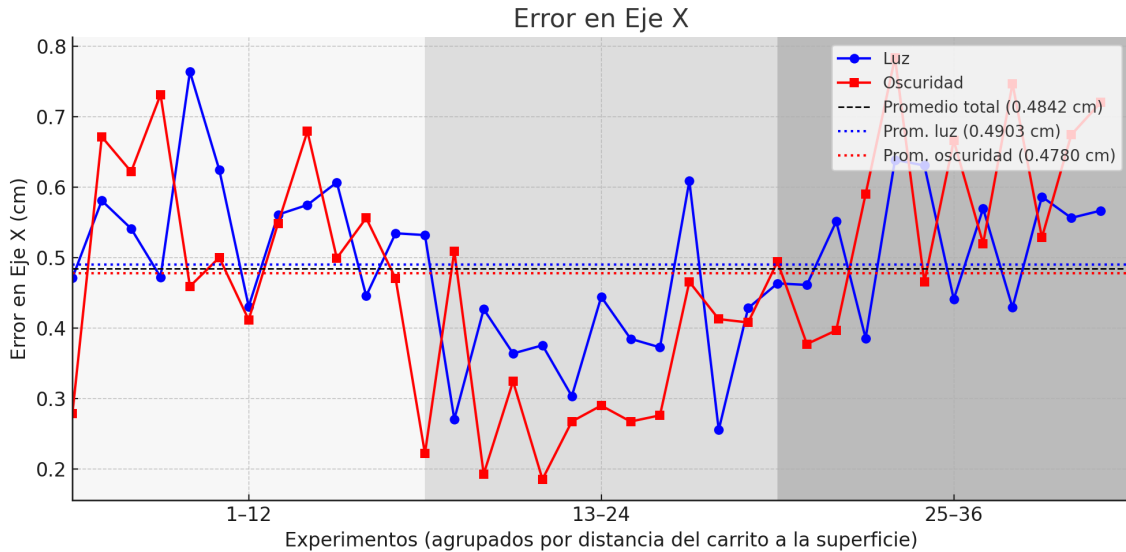


Figura 4.30: Error de correspondencia en el eje X para luz ambiental y oscuridad, representando en cada tono de gris una distancia a la superficie: 20 cm, 30 cm y 40 cm.

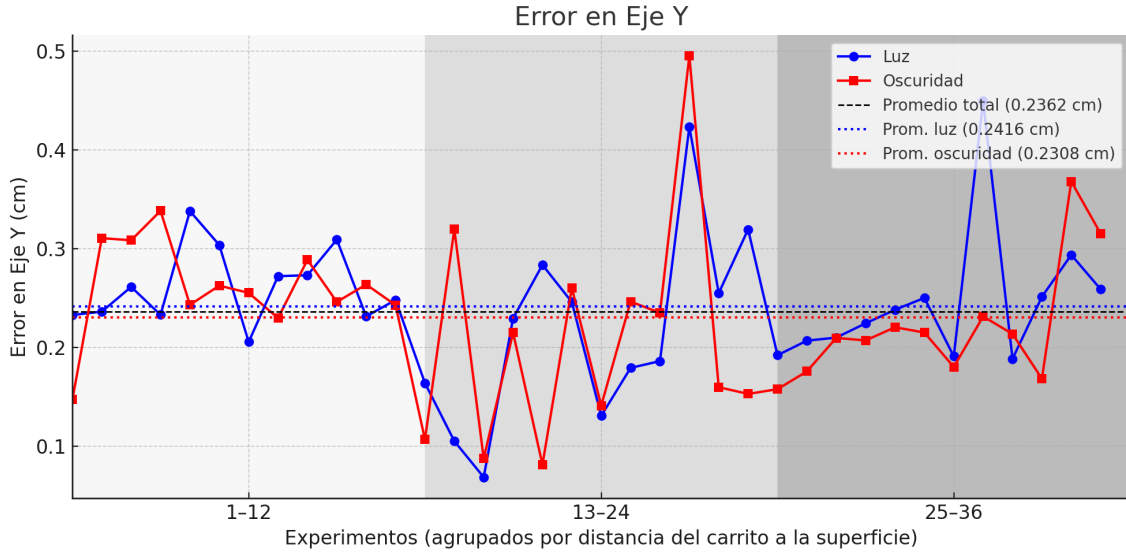


Figura 4.31: Error de correspondencia en el eje Y para luz ambiental y oscuridad, representando en cada tono de gris una distancia a la superficie: 20 cm, 30 cm y 40 cm.

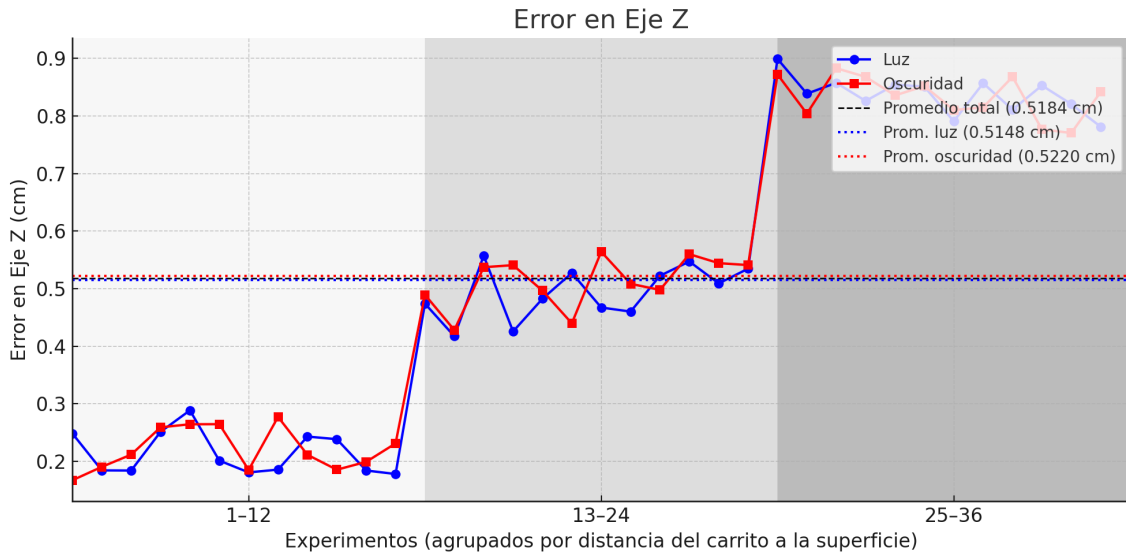


Figura 4.32: Error de correspondencia en el eje Z para luz ambiental y oscuridad, representando en cada tono de gris una distancia a la superficie: 20 cm, 30 cm y 40 cm.

Como en el caso de la calibración realizada sobre mesa óptica, el error se define como la discrepancia entre los puntos homólogos tras aplicar la transformación geométrica. La Tabla 4.5 presenta la media y la desviación estándar del error en los tres ejes cartesianos, calculadas a partir de las 72 mediciones realizadas (36 con cada nivel de iluminación).

En condiciones ideales, ambos sistemas deberían reportar coordenadas coincidentes para cada pun-

to después de la transformación. El valor promedio del error cuantifica la magnitud media de esta discrepancia, mientras que la desviación estándar caracteriza la dispersión de los errores individuales. Estas métricas permiten evaluar la precisión y la estabilidad del método propuesto a lo largo de múltiples repeticiones experimentales, en concordancia con los principios de trazabilidad y repetibilidad de la metrología dimensional [67].

Eje	Error promedio (cm)	Desviación estándar (cm)
X	0.4842	0.1419
Y	0.2362	0.07804
Z	0.5184	0.2571

Tabla 4.5: Resumen de resultados de calibración extrínseca considerando 72 experimentos sobre robot con movimiento intermitente.

Los resultados globales indican que el método mantiene un desempeño robusto bajo distintas condiciones de iluminación. Por ejemplo, el error promedio en el eje X fue de 0.4903 cm con luz ambiental y 0.4780 cm en oscuridad, con una ligera diferencia en la dispersión. En el eje Z , incluso en ausencia de luz, la variación fue mínima, lo que refleja una alta estabilidad del sistema.

También se identificó que el incremento en la distancia introduce mayor error y dispersión, especialmente en el eje Z , donde se alcanzó un valor promedio de 0.8349 cm a 40 cm, comparado con apenas 0.2172 cm a 20 cm. Esta tendencia sugiere una mayor sensibilidad del sistema a la profundidad conforme aumenta la separación entre sensores y superficie.

Por otra parte, el análisis del número de posiciones k empleadas en cada barrido revela que el método propuesto mantiene un desempeño consistente incluso con valores bajos de k . En particular, los errores promedio en los tres ejes permanecen en rangos comparables entre $k = 3$ y $k = 6$, sin evidenciar un incremento significativo al reducir la cantidad de puntos muestreados. Esta estabilidad sugiere que el procedimiento es capaz de generar transformaciones confiables incluso con conjuntos de datos reducidos, lo cual representa una ventaja importante en contextos donde se requiere minimizar el tiempo de captura o el consumo computacional. Además, esta propiedad refuerza la aplicabilidad del método en entornos dinámicos o de recursos limitados, sin comprometer la precisión de la calibración extrínseca.

Nuevamente, el conjunto de los resultados obtenidos permiten afirmar que el método de calibración extrínseca propuesto conserva su precisión y estabilidad incluso en escenarios con condiciones no controladas, como iluminación reducida, mayor distancia sensor-superficie y adquisición limitada de datos. Esta validación sobre un sistema móvil con movimiento intermitente constituye un avance significativo hacia su implementación en contextos operativos reales. A partir de esta base experimental, se abre la posibilidad de extender la metodología a escenarios de movimiento continuo, en los cuales la estimación de la transformación extrínseca deberá realizarse dinámicamente durante el desplazamiento del sistema, lo que representa una línea prometedora para investigaciones futuras.

Capítulo 5: Conclusiones

Este trabajo presentó un método sistemático de calibración conjunta para un sistema híbrido de percepción tridimensional compuesto por un Sistema de Visión Técnica (TVS) y un sistema de estereovisión, capaz de operar sin patrones de referencia impresos, algoritmos de optimización iterativa ni sensores adicionales. La propuesta abarcó tanto la calibración intrínseca como la extrínseca, integrando estrategias experimentales y computacionales que permitieron una alineación geométrica robusta, reproducible y adaptable a distintos entornos operativos.

Para el sistema estéreo, se aplicó el método de Zhang con el objetivo de estimar los parámetros intrínsecos de ambas cámaras, obteniendo las matrices necesarias para la reconstrucción tridimensional: la matriz de rotación \mathbf{R} , el vector de traslación \mathbf{T} . Posteriormente se determinó la matriz de reproyección \mathbf{Q} del sistema estéreo, el cual se utilizó no solo para reconstrucción, sino como herramienta auxiliar en el proceso de caracterización geométrica del TVS.

En relación con los objetivos planteados, todos fueron alcanzados de forma satisfactoria. Se determinó con precisión la calibración intrínseca del sistema estéreo mediante el método de Zhang, asegurando una reconstrucción 3D confiable. La estereovisión se utilizó con éxito como herramienta de caracterización geométrica del TVS, permitiendo estimar parámetros intrínsecos del propio sistema, relevantes para el cálculo de coordenadas mediante triangulación dinámica. Se desarrolló un procedimiento automatizado para la calibración extrínseca sin necesidad de patrones impresos, basado en detección láser y ajuste de planos mediante SVD, que demostró ser funcional tanto en laboratorio como en un robot con movimiento intermitente. Asimismo, se verificó la viabilidad del método en condiciones de baja iluminación, destacando su utilidad en entornos confinados. También se avanzó en la integración de los procedimientos de calibración en la interfaz operativa del TVS, facilitando su ejecución por parte del usuario final.

Desde el punto de vista de calibración intrínseca, se caracterizó experimentalmente el comportamiento angular del TVS mediante el uso de la estereovisión como herramienta de medición. Esto

permitió estimar con precisión la relación grados/paso de los motores y establecer un procedimiento cuantitativo para evaluar el backlash mecánico en los cuatro sentidos de movimiento del posicionador. La triangulación estereoscópica aportó robustez frente a errores de colocación manual de la superficie de referencia, lo cual simplifica la implementación y mejora la portabilidad del método. Además, se logró una adecuada coalineación de los espejos que integran el TVS, optimizando así la calidad y estabilidad del haz láser proyectado. Los resultados de la coalineación pueden mejorar si se modifica el diseño del prototipo para dar más libertad a ajustes mecánicos.

En condiciones de laboratorio, la calibración extrínseca se realizó utilizando una superficie plana sin marcas visibles, generando nubes de puntos correspondientes en ambos subsistemas. Se propuso un esquema de alineación basado en la estimación de transformaciones rígidas: la rotación se determinó mediante el ajuste de planos por descomposición en valores singulares (SVD, por sus siglas en inglés) y la alineación de vectores normales asociados a cada plano, mientras que la traslación se estimó a partir de la correspondencia punto a punto entre los centroides de las nubes. El procedimiento utilizando una mesa óptica alcanzó errores promedio de:

$$x_e = 0.2741 \text{ cm}, \quad y_e = 0.8843 \text{ cm}, \quad z_e = 1.5090 \text{ cm},$$

los cuales representan la distancia media entre coordenadas correspondientes de dos nubes de puntos. Valores bajos de error permiten una integración efectiva de información tridimensional entre los subsistemas, lo cual es útil en aplicaciones como la planificación de trayectorias o la detección estructural de objetos. Por otro lado, el método no impone restricciones estrictas sobre la posición del plano de referencia, lo cual otorga versatilidad espacial y escalabilidad.

Posteriormente, se validó la metodología en un entorno móvil donde el sistema híbrido fue montado sobre un robot que realizaba desplazamientos intermitentes frente a una pared plana. La calibración extrínseca conservó su precisión, con errores medios de correspondencia de:

$$\varepsilon_x = 0.4842 \text{ cm}, \quad \varepsilon_y = 0.2362 \text{ cm}, \quad \varepsilon_z = 0.5184 \text{ cm}.$$

Además, se comprobó su funcionalidad bajo condiciones de baja iluminación, gracias al uso del láser del TVS como referencia activa, lo cual lo hace especialmente adecuado para aplicaciones en interiores oscuros, como tuberías, túneles o ductos de ventilación.

En general, el método propuesto demuestra una alta resiliencia frente a condiciones no ideales, tales como la ausencia de patrones impresos, oclusiones parciales, iluminación deficiente y variaciones en la distancia u orientación de las superficies observadas. No obstante, es importante reconocer ciertas limitaciones técnicas. La precisión del ajuste extrínseco depende de la calidad geométrica de la superficie de referencia: desviaciones respecto a la planaridad ideal pueden introducir errores sistemáticos en la estimación de los planos y de sus vectores normales asociados. Asimismo, la visibilidad y nitidez de la mancha láser influyen directamente en la detección confiable de sus centroides, siendo susceptibles a degradación por efectos como reflexiones especulares o absorción superficial elevada.

En síntesis, la metodología desarrollada constituye una solución precisa, portátil y libre de dependencias rígidas para la calibración de sistemas multisensoriales. Su eficacia se comprobó tanto en ambientes controlados como móviles, sentando las bases para aplicaciones avanzadas en reconstrucción 3D, navegación autónoma e inspección estructural en escenarios adversos. Como líneas futuras de desarrollo, se plantea la generalización del procedimiento para incluir superficies no planas o con geometría irregular, así como su implementación en sistemas con movilidad continua o capacidad de operación en tiempo real. Adicionalmente, se propone incorporar mecanismos de calibración en línea que operen sin detener el sistema y validar el método en escenarios reales de inspección estructural en espacios confinados y con escasa iluminación. Finalmente, se vislumbra la posibilidad de prescindir del método de Zhang mediante un enfoque auto-informado, en el que el TVS proyecta manchas láser que generan correspondencias 3D-2D en el sistema estéreo, permitiendo ejecutar un ciclo iterativo de calibración conjunta.

Capítulo 6: Apéndice A: Publicaciones

A continuación se presenta el listado de publicaciones científicas generadas durante el desarrollo de esta tesis:

Artículos de revista:

- Extrinsic calibration of complex machine vision system for mobile robot.

Integration

- Modeling and Analysis of Nonlinear Chaotic Mechanical Dynamics in Laser Scanning Systems.

Chaos Theory and Applications

Artículos de conferencia:

- Joint calibration of Machine Vision subsystems for robust surrounding 3D perception.

IECON 2024

- Robust Control Strategy for Laser Scanning Systems with Friction and Mechanical Vibration Compensation.

IECON 2024

- Integration of laser scanning and projection speckle pattern for advanced pipeline monitoring.

IECON 2024

Bibliografía

- [1] K. Bimbraw, “Autonomous cars: Past, present and future—A review of the developments in the last century, the present scenario and the expected future of autonomous vehicle technology,” en *2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2015, pp. 191-198.
- [2] A. Dawda y M. Nguyen, “Defects detection in highly specular surface using a combination of stereo and laser reconstruction,” en *2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 2020, pp. 1-6. DOI: 10.1109/IVCNZ51579.2020.9290660.
- [3] C. Sepulveda-Valdez, O. Sergiyenko, V. Tyrsa et al., “Mathematical modeling for robot 3D laser scanning in complete darkness environments to advance pipeline inspection,” *Mathematics*, vol. 12, n.º 13, 2024. DOI: 10.3390/math12131940.
- [4] Z. Yang y A. Ganz, “Vision-based responders localization techniques in urban search and rescue scenarios,” en *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 2640-2643. DOI: 10.1109/EMBC.2016.7591272.
- [5] R. Wang, “3D reconstruction using a linear laser scanner and a camera,” en *2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE)*, 2021, pp. 666-673. DOI: 10.1109/ICAICE54393.2021.00131.
- [6] M. N. A. Wahab, N. Sivadev y K. Sundaraj, “Target distance estimation using monocular vision system for mobile robot,” en *2011 IEEE Conference on Open Systems*, 2011, pp. 11-15. DOI: 10.1109/ICOS.2011.6079296.
- [7] H. Lin, H. Zhang, J. Huo, J. Li, H. Zhang e Y. Li, “High-precision 3D reconstruction of underwater concrete using integrated line structured light and stereo vision,” *Automation in Construction*, vol. 169, pág. 105 883, 2025. DOI: 10.1016/j.autcon.2024.105883.
- [8] D. Scaramuzza, “Omnidirectional Camera,” en *Computer Vision: A Reference Guide*, K. Ikeuchi, ed. Boston, MA: Springer US, 2014, pp. 552-560, ISBN: 978-0-387-31439-6. DOI: 10.1007/978-0-387-31439-6_488. dirección: https://doi.org/10.1007/978-0-387-31439-6_488.
- [9] J. Villacrés y S. Vougioukas, “Assessing a multi-camera system to enhance fruit visibility for robotic harvesting in a V-trellised apple orchard,” *Computers and Electronics in Agriculture*, vol. 224, pág. 109 164, 2024. DOI: 10.1016/j.compag.2024.109164.
- [10] C. Sepulveda-Valdez, O. Sergiyenko, R. Alaniz-Plata et al., “Laser scanning point cloud improvement by implementation of RANSAC for pipeline inspection application,” en *IECON 2023 - 49th Annual Conference of the IEEE Industrial Electronics Society*, 2023, pp. 1-6. DOI: 10.1109/IECON51785.2023.10312684.
- [11] K. V. Mack, W. D. Jemison, L. K. Rumbaugh, D. W. Illig y M. K. Banavar, “Time-of-flight (ToF) cameras for underwater situational awareness,” en *OCEANS 2019 - Seattle*, IEEE, 2019. DOI: 10.1109/OCEANS40490.2019.8962555.
- [12] L. C. Básaca-Preciado, O. Y. Sergiyenko, J. C. Rodríguez-Quinonez et al., “Optical 3D laser measurement system for navigation of autonomous mobile robot,” *Optics and Lasers in Engineering*, vol. 54, pp. 159-169, 2014. DOI: 10.1016/j.optlaseng.2013.08.005.
- [13] Y. Ding, J. Liu, J. Ye, W. Xiang, H.-C. Wu y C. Busch, “3D LiDAR and color camera data fusion,” en *2020 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2020, pp. 1-4. DOI: 10.1109/BMSB49480.2020.9379430.
- [14] R. Szeliski, *Computer vision: Algorithms and applications* (Texts in Computer Science), 2.^a ed. Cham: Springer Nature, 2022, ISBN: 978-3-030-34371-2. DOI: 10.1007/978-3-030-34372-9. dirección: <https://link.springer.com/book/10.1007/978-3-030-34372-9>.

- [15] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, n.º 11, pp. 1330-1334, 2000. DOI: 10.1109/34.888718.
- [16] X. Garcia-Cruz, O. Sergiyenko, V. Tyrsa et al., "Optimization of 3D laser scanning speed by use of combined variable step," *Optics and Lasers in Engineering*, vol. 54, pp. 141-151, 2014, ISSN: 0143-8166. DOI: <https://doi.org/10.1016/j.optlaseng.2013.08.011>. dirección: <https://www.sciencedirect.com/science/article/pii/S0143816613002546>.
- [17] O. Y. Sergiyenko y V. V. Tyrsa, "3D optical machine vision sensors with intelligent data management for robotic swarm navigation improvement," *IEEE Sensors Journal*, vol. 21, n.º 10, pp. 11 262-11 274, 2021. DOI: 10.1109/JSEN.2020.3007856.
- [18] O. Sergiyenko, V. Tyrsa, A. Zhirabok y A. Zuev, "Sliding mode observer based fault identification in automatic vision system of robot," *Control Engineering Practice*, vol. 139, pág. 105 614, 2023. DOI: 10.1016/j.conengprac.2023.105614.
- [19] J. C. Rodríguez-Quiñonez, O. Y. Sergiyenko, L. C. B. Preciado et al., "Optical monitoring of scoliosis by 3D medical laser scanner," *Optics and Lasers in Engineering*, vol. 54, pp. 175-186, 2014, ISSN: 0143-8166. DOI: <https://doi.org/10.1016/j.optlaseng.2013.07.026>. dirección: <https://www.sciencedirect.com/science/article/pii/S014381661300242X>.
- [20] O. Sergiyenko, R. Alaniz-Plata, W. Flores-Fuentes et al., "Multi-view 3D data fusion and patching to reduce Shannon entropy in robotic vision," *Optics and Lasers in Engineering*, vol. 177, 2024. DOI: 10.1016/j.optlaseng.2024.108132.
- [21] A. Khurana y K. S. Nagla, "Extrinsic calibration methods for laser range finder and camera: A systematic review," *MAPAN*, vol. 36, n.º 3, pp. 669-690, 2021. DOI: 10.1007/s12647-021-00500-x.
- [22] Q. Zhang y R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," en *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, 2004, pp. 2301-2306. DOI: 10.1109/IROS.2004.1389752.
- [23] C. Mei y P. Rives, "Calibration between a central catadioptric camera and a laser range finder for robotic applications," en *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2006, pp. 532-537. DOI: 10.1109/ROBOT.2006.1641765.
- [24] C. D. Lombard y C. E. van Daalen, "Extrinsic calibration of a push-broom lidar and camera using 3-D multi-planar association," en *Proceedings of the PRASA-RobMech International Conference*, 2016, pp. 1-6. DOI: 10.1109/RoboMech.2016.7813180.
- [25] G. Li, Y. Liu, L. Dong, X. Cai y D. Zhou, "An algorithm for extrinsic parameters calibration of a camera and a laser range finder using line features," en *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2007, pp. 3854-3859. DOI: 10.1109/IROS.2007.4399041.
- [26] L. Zhou, "A new minimal solution for the extrinsic calibration of a 2D LIDAR and a camera using three plane-line correspondences," *IEEE Sensors Journal*, vol. 14, n.º 2, pp. 442-454, 2014. DOI: 10.1109/JSEN.2013.2284789.
- [27] Z. Hu, Y. Li, N. Li y B. Zhao, "Extrinsic calibration of 2-D laser rangefinder and camera from single shot based on minimal solution," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, n.º 4, pp. 915-929, 2016. DOI: 10.1109/TIM.2016.2518248.
- [28] S. Mishra, G. Pandey y S. Saripalli, "Extrinsic calibration of a 3D-LIDAR and a camera," *CoRR*, vol. abs/2003.01213, 2020. DOI: 10.48550/arXiv.2003.01213. eprint: 2003.01213.
- [29] P. Núñez, P. Drews-Jr, R. Rocha y J. Dias, "Data Fusion Calibration for a 3D Laser Range Finder and a Camera using Inertial Data," ene. de 2009, pp. 31-36.
- [30] H. Aliakbarpour, P. Nunez, J. Prado, K. Khoshhal y J. Dias, "An efficient algorithm for extrinsic calibration between a 3D laser range finder and a stereo camera for surveillance," en *Proceedings of the International Conference on Advanced Robotics*, 2009, pp. 1-6.

- [31] S. Rodriguez, V. Fremont y P. Bonnifait, “Extrinsic calibration between a multi-layer lidar and a camera,” en *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2008, pp. 214-219. DOI: 10.1109/MFI.2008.4648067.
- [32] H. Xu, G. Lan, S. Wu y Q. Hao, “Online intelligent calibration of cameras and LiDARs for autonomous driving systems,” en *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 3913-3920. DOI: 10.1109/ITSC.2019.8916872.
- [33] M. Nixon y A. Aguado, *Feature Extraction and Image Processing for Computer Vision*, 4.^a ed. Oxford: Academic Press, 2019, ISBN: 978-0-12-814976-8.
- [34] “Multiview Geometry,” en *Computer Vision: A Reference Guide*, K. Ikeuchi, ed. Boston, MA: Springer US, 2014, pp. 516-516, ISBN: 978-0-387-31439-6. DOI: 10.1007/978-0-387-31439-6_100011. dirección: https://doi.org/10.1007/978-0-387-31439-6_100011.
- [35] X. Lian, Z. Liu, X. Wang y L. Dou, “Reconstructing indoor environmental 3D model using laser range scanners and omnidirectional camera,” *Proceedings of the 2008 World Congress on Intelligent Control and Automation (WCICA)*, 2008. DOI: 10.1109/WCICA.2008.4594462.
- [36] H.-L. Tsay, C.-H. Lin, M.-F. Chen, P.-M. Lin, C.-C. Lien y H.-J. Kuo, “Radiometric and geometric correction methods for SWIR cameras,” en *2022 25th International Conference on Mechatronics Technology (ICMT)*, 2022, pp. 1-5. DOI: 10.1109/ICMT56556.2022.9997793.
- [37] M. Kobayashi, K. Yamazaki, F. G. Caballero et al., “Wide angle multi-shift stereo camera with monocular vision,” en *2020 IEEE International Conference on Consumer Electronics (ICCE)*, 2020, pp. 1-4. DOI: 10.1109/ICCE46568.2020.9043079.
- [38] S. Gidel, C. Blanc, T. Chateau, P. Checchin y L. Trassoudaine, “Non-parametric laser and video data fusion: Application to pedestrian detection in urban environment,” en *2009 12th International Conference on Information Fusion*, 2009, pp. 626-632.
- [39] Z. Zhang, “Camera calibration,” en *Computer Vision: A Reference Guide*, K. Ikeuchi, ed. Boston, MA: Springer US, 2014, pp. 76-77, ISBN: 978-0-387-31439-6. DOI: 10.1007/978-0-387-31439-6_164. dirección: https://doi.org/10.1007/978-0-387-31439-6_164.
- [40] T. Dieterle, F. Particke, L. Patino-Studencki y J. Thielecke, “Sensor data fusion of LiDAR with stereo RGB-D camera for object tracking,” en *2017 IEEE SENSORS*, 2017, pp. 1-3. DOI: 10.1109/ICSENS.2017.8234267.
- [41] W. Xu, A. Bosseboeuf, F. Parrain y E. Martincic, “Design of a long range bidirectional MEMS scanner for a tunable 3D integrated Mirau interferometer,” en *2014 Symposium on Design, Test, Integration and Packaging of MEMS/MOEMS (DTIP)*, 2014, pp. 1-6. DOI: 10.1109/DTIP.2014.7056655.
- [42] L. C. Básaca, J. Rodríguez, O. Y. Sergiyenko et al., “3D laser scanning vision system for autonomous robot navigation,” en *2010 IEEE International Symposium on Industrial Electronics*, 2010, pp. 1773-1778. DOI: 10.1109/ISIE.2010.5637874.
- [43] C. S. Silva y P. Wimalaratne, “Sensor fusion for visually impaired navigation in constrained spaces,” en *2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS)*, 2016, pp. 1-6. DOI: 10.1109/ICIAfS.2016.7946537.
- [44] J. Wang, C. Zhang, W. Zhu, Z. Zhang, Z. Xiong y P. A. Chou, “3D scene reconstruction by multiple structured-light based commodity depth cameras,” en *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 5429-5432. DOI: 10.1109/ICASSP.2012.6289149.
- [45] N. El-Sheimy, K.-W. Chiang y A. Noureldin, “The utilization of artificial neural networks for multisensor system integration in navigation and positioning instruments,” *IEEE Transactions on Instrumentation and Measurement*, vol. 55, n.º 5, pp. 1606-1615, 2006. DOI: 10.1109/TIM.2006.881033.
- [46] C. H. Lehmann, *Geometría Analítica*. México: Editorial Limusa S.A. de C.V., 1984, ISBN: 968-18-1176-3.

- [47] B. Li, M. Liu, X. Wang y X. Li, “An adaptive plug-and-play online extrinsic calibration approach for LiDAR-camera fusion,” *Measurement*, vol. 248, pág. 116877, 2025. DOI: 10.1016/j.measurement.2025.116877. dirección: <https://www.sciencedirect.com/science/article/pii/S0263224125002362>.
- [48] ISO/IEC, *International vocabulary of metrology – Basic and general concepts and associated terms (VIM)*, 3.^a ed. Geneva: International Organization for Standardization, 2007, ISO/IEC Guide 99:2007.
- [49] T. Hanning, *High Precision Camera Calibration*. Vieweg+Teubner Verlag Wiesbaden, 2011, ISBN: 978-3-8348-1413-5. DOI: 10.1007/978-3-8348-9830-2. dirección: <https://doi.org/10.1007/978-3-8348-9830-2>.
- [50] Z. Tang, “High precision camera calibration,” HAL ID: tel-00675484, Tesis doct., École Normale Supérieure de Cachan - ENS Cachan, 2011. dirección: <https://theses.hal.science/tel-00675484>.
- [51] P. Sturm, “Pinhole camera model,” en *Computer Vision: A Reference Guide*, K. Ikeuchi, ed. Boston, MA: Springer US, 2014, pp. 610-613, ISBN: 978-0-387-31439-6. DOI: 10.1007/978-0-387-31439-6_472. dirección: https://doi.org/10.1007/978-0-387-31439-6_472.
- [52] L. Wang y G. Liu, “Three camera lens distortion correction models and its application,” en *2022 3rd International Conference on Geology, Mapping and Remote Sensing (ICGMRS)*, 2022, pp. 462-467. DOI: 10.1109/ICGMRS55602.2022.9849271.
- [53] J. Zhao, D.-M. Yan, G.-Z. Men e Y.-K. Zhang, “A method of calibrating the intrinsic and extrinsic camera parameters separately for multi-camera systems,” en *2007 International Conference on Machine Learning and Cybernetics*, vol. 3, 2007, pp. 1548-1553. DOI: 10.1109/ICMLC.2007.4370391.
- [54] F.-Y. Wang, “An efficient coordinate frame calibration method for 3-D measurement by multiple camera systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 35, n.º 4, pp. 453-464, 2005. DOI: 10.1109/TSMCC.2004.843208.
- [55] O. Kainz, F. Jakab, P. Fecil’ak, R. Vápeník, A. Deák y D. Cymbalák, “Estimation of camera intrinsic matrix parameters and its utilization in the extraction of dimensional units,” en *2016 International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 2016, pp. 153-156. DOI: 10.1109/ICETA.2016.7802057.
- [56] K. El-Ashmawy, “Using direct linear transformation (DLT) method for aerial photogrammetry applications,” *Geodesy and Cartography*, vol. 44, pp. 71-79, 2018. DOI: 10.3846/gac.2018.1629.
- [57] L. Wu y B. Zhu, “Binocular stereovision camera calibration,” en *2015 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2015, pp. 2638-2642. DOI: 10.1109/ICMA.2015.7237903.
- [58] R. Tsai, “A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses,” *IEEE Journal on Robotics and Automation*, vol. 3, n.º 4, pp. 323-344, 1987. DOI: 10.1109/JRA.1987.1087109.
- [59] R. Koch y J. Bruenger, “Depth estimation,” en *Computer Vision: A Reference Guide*, K. Ikeuchi, ed. Cham: Springer International Publishing, 2021, pp. 290-294, ISBN: 978-3-030-63416-2. DOI: 10.1007/978-3-030-63416-2_125. dirección: https://doi.org/10.1007/978-3-030-63416-2_125.
- [60] Y. Cui, F. Zhou, Y. Wang, L. Liu y H. Gao, “Precise calibration of binocular vision system used for vision measurement,” *Opt. Express*, vol. 22, n.º 8, pp. 9134-9149, 2014. DOI: 10.1364/OE.22.009134. dirección: <https://opg.optica.org/oe/abstract.cfm?URI=oe-22-8-9134>.
- [61] S. Tian, L. Shan y Z. Liu, “Levenberg-Marquardt algorithm based nonlinear optimization of camera calibration for relative measurement,” en *2015 34th Chinese Control Conference (CCC)*, 2015, pp. 4868-4872. DOI: 10.1109/ChiCC.2015.7260394.

- [62] M. A. Goshtasby, "Finding homologous points," en *Theory and Applications of Image Registration*. John Wiley & Sons, Ltd, 2017, cap. 5, pp. 133-214, ISBN: 9781119171744. DOI: 10.1002/9781119171744.ch5. dirección: <https://doi.org/10.1002/9781119171744.ch5>.
- [63] R. A. Hamzah, A. M. A. Hamid y S. I. M. Salim, "The solution of stereo correspondence problem using block matching algorithm in stereo vision mobile robot," en *2010 Second International Conference on Computer Research and Development*, 2010, pp. 733-737. DOI: 10.1109/ICCRD.2010.167.
- [64] G. Li, "Stereo matching using normalized cross-correlation in LogRGB space," en *2012 International Conference on Computer Vision in Remote Sensing*, 2012, pp. 19-23. DOI: 10.1109/CVRS.2012.6421226.
- [65] W. Wang y H. T. Tsui, "A SVD decomposition of essential matrix with eight solutions for the relative positions of two perspective cameras," en *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, vol. 1, 2000, pp. 362-365. DOI: 10.1109/ICPR.2000.905353.
- [66] MathWorks, *Prepare Camera and Capture Images For Camera Calibration*, Versión consultada el 19 de abril de 2025, MathWorks, 2025. dirección: <https://www.mathworks.com/help/vision/ug/prepare-camera-and-capture-images-for-camera-calibration.html>.
- [67] A. A. Hawam, "The quantifying of uncertainty in measurement," en *Handbook of Metrology and Applications*, D. K. Aswal, S. Yadav, T. Takatsuji, P. Rachakonda y H. Kumar, eds. Singapore: Springer Nature Singapore, 2023, pp. 2377-2407, ISBN: 978-981-99-2074-7. DOI: 10.1007/978-981-99-2074-7_126. dirección: https://doi.org/10.1007/978-981-99-2074-7_126.